PRINCETON LECTURES IN ANALYSIS I

FOURIER ANALYSIS

AN INTRODUCTION



ELIAS M. STEIN & RAMI SHAKARCHI

FOURIER ANALYSIS

Princeton Lectures in Analysis

I Fourier Analysis: An Introduction

II Complex Analysis

III Real Analysis:
Measure Theory, Integration, and
Hilbert Spaces

PRINCETON LECTURES IN ANALYSIS

I

FOURIER ANALYSIS

AN INTRODUCTION

Elias M. Stein

&

Rami Shakarchi

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

Library of Congress Control Number 2003103688
ISBN 978-0-691-11384-5

British Library Cataloging-in-Publication Data is available

The publisher would like to acknowledge the authors of this volume for
providing the camera-ready copy from which this book was printed

Printed on acid-free paper. ∞

press.princeton.edu

Printed in the United States of America

5 7 9 10 8 6

To my grandchildren
Carolyn, Alison, Jason

E.M.S.

To my parents
Mohamed & Mireille
and my brother
Karim

R.S.

Foreword

Beginning in the spring of 2000, a series of four one-semester courses were taught at Princeton University whose purpose was to present, in an integrated manner, the core areas of analysis. The objective was to make plain the organic unity that exists between the various parts of the subject, and to illustrate the wide applicability of ideas of analysis to other fields of mathematics and science. The present series of books is an elaboration of the lectures that were given.

While there are a number of excellent texts dealing with individual parts of what we cover, our exposition aims at a different goal: presenting the various sub-areas of analysis not as separate disciplines, but rather as highly interconnected. It is our view that seeing these relations and their resulting synergies will motivate the reader to attain a better understanding of the subject as a whole. With this outcome in mind, we have concentrated on the main ideas and theorems that have shaped the field (sometimes sacrificing a more systematic approach), and we have been sensitive to the historical order in which the logic of the subject developed.

We have organized our exposition into four volumes, each reflecting the material covered in a semester. Their contents may be broadly summarized as follows:

I. Fourier series and integrals.

II. Complex analysis.

III. Measure theory, Lebesgue integration, and Hilbert spaces.

IV. A selection of further topics, including functional analysis, distributions, and elements of probability theory.

However, this listing does not by itself give a complete picture of the many interconnections that are presented, nor of the applications to other branches that are highlighted. To give a few examples: the elements of (finite) Fourier series studied in Book I, which lead to Dirichlet characters, and from there to the infinitude of primes in an arithmetic progression; the X-ray and Radon transforms, which arise in a number of

problems in Book I, and reappear in Book III to play an important role in understanding Besicovitch-like sets in two and three dimensions; Fatou's theorem, which guarantees the existence of boundary values of bounded holomorphic functions in the disc, and whose proof relies on ideas developed in each of the first three books; and the theta function, which first occurs in Book I in the solution of the heat equation, and is then used in Book II to find the number of ways an integer can be represented as the sum of two or four squares, and in the analytic continuation of the zeta function.

A few further words about the books and the courses on which they were based. These courses where given at a rather intensive pace, with 48 lecture-hours a semester. The weekly problem sets played an indispensable part, and as a result exercises and problems have a similarly important role in our books. Each chapter has a series of "Exercises" that are tied directly to the text, and while some are easy, others may require more effort. However, the substantial number of hints that are given should enable the reader to attack most exercises. There are also more involved and challenging "Problems"; the ones that are most difficult, or go beyond the scope of the text, are marked with an asterisk.

Despite the substantial connections that exist between the different volumes, enough overlapping material has been provided so that each of the first three books requires only minimal prerequisites: acquaintance with elementary topics in analysis such as limits, series, differentiable functions, and Riemann integration, together with some exposure to linear algebra. This makes these books accessible to students interested in such diverse disciplines as mathematics, physics, engineering, and finance, at both the undergraduate and graduate level.

It is with great pleasure that we express our appreciation to all who have aided in this enterprise. We are particularly grateful to the students who participated in the four courses. Their continuing interest, enthusiasm, and dedication provided the encouragement that made this project possible. We also wish to thank Adrian Banner and Jose Luis Rodrigo for their special help in running the courses, and their efforts to see that the students got the most from each class. In addition, Adrian Banner also made valuable suggestions that are incorporated in the text.

We wish also to record a note of special thanks for the following individuals: Charles Fefferman, who taught the first week (successfully launching the whole project!); Paul Hagelstein, who in addition to reading part of the manuscript taught several weeks of one of the courses, and has since taken over the teaching of the second round of the series; and Daniel Levine, who gave valuable help in proof-reading. Last but not least, our thanks go to Gerree Pecht, for her consummate skill in typesetting and for the time and energy she spent in the preparation of all aspects of the lectures, such as transparencies, notes, and the manuscript.

We are also happy to acknowledge our indebtedness for the support we received from the 250th Anniversary Fund of Princeton University, and the National Science Foundation's VIGRE program.

Elias M. Stein

Rami Shakarchi

Princeton, New Jersey
August 2002

Preface to Book I

Any effort to present an overall view of analysis must at its start deal with the following questions: Where does one begin? What are the initial subjects to be treated, and in what order are the relevant concepts and basic techniques to be developed?

Our answers to these questions are guided by our view of the centrality of Fourier analysis, both in the role it has played in the development of the subject, and in the fact that its ideas permeate much of the present-day analysis. For these reasons we have devoted this first volume to an exposition of some basic facts about Fourier series, taken together with a study of elements of Fourier transforms and finite Fourier analysis. Starting this way allows one to see rather easily certain applications to other sciences, together with the link to such topics as partial differential equations and number theory. In later volumes several of these connections will be taken up from a more systematic point of view, and the ties that exist with complex analysis, real analysis, Hilbert space theory, and other areas will be explored further.

In the same spirit, we have been mindful not to overburden the beginning student with some of the difficulties that are inherent in the subject: a proper appreciation of the subtleties and technical complications that arise can come only after one has mastered some of the initial ideas involved. This point of view has led us to the following choice of material in the present volume:

- Fourier series. At this early stage it is not appropriate to introduce measure theory and Lebesgue integration. For this reason our treatment of Fourier series in the first four chapters is carried out in the context of Riemann integrable functions. Even with this restriction, a substantial part of the theory can be developed, detailing convergence and summability; also, a variety of connections with other problems in mathematics can be illustrated.

- Fourier transform. For the same reasons, instead of undertaking the theory in a general setting, we confine ourselves in Chapters 5 and 6 largely to the framework of test functions. Despite these limitations, we can learn a number of basic and interesting facts about Fourier analysis in $\mathbb{R}^d$ and its relation to other areas, including the wave equation and the Radon transform.

- Finite Fourier analysis. This is an introductory subject *par excellence*, because limits and integrals are not explicitly present. Nevertheless, the subject has several striking applications, including the proof of the infinitude of primes in arithmetic progression.

Taking into account the introductory nature of this first volume, we have kept the prerequisites to a minimum. Although we suppose some acquaintance with the notion of the Riemann integral, we provide an appendix that contains most of the results about integration needed in the text.

We hope that this approach will facilitate the goal that we have set for ourselves: to inspire the interested reader to learn more about this fascinating subject, and to discover how Fourier analysis affects decisively other parts of mathematics and science.

Contents

1 The Genesis of Fourier Analysis

> Regarding the researches of d'Alembert and Euler could one not add that if they knew this expansion, they made but a very imperfect use of it. They were both persuaded that an arbitrary and discontinuous function could never be resolved in series of this kind, and it does not even seem that anyone had developed a constant in cosines of multiple arcs, the first problem which I had to solve in the theory of heat.
>
> *J. Fourier,* 1808-9

In the beginning, it was the problem of the vibrating string, and the later investigation of heat flow, that led to the development of Fourier analysis. The laws governing these distinct physical phenomena were expressed by two different partial differential equations, the wave and heat equations, and these were solved in terms of Fourier series.

Here we want to start by describing in some detail the development of these ideas. We will do this initially in the context of the problem of the vibrating string, and we will proceed in three steps. First, we describe several physical (empirical) concepts which motivate corresponding mathematical ideas of importance for our study. These are: the role of the functions $\cos t$, $\sin t$, and e^{it} suggested by simple harmonic motion; the use of separation of variables, derived from the phenomenon of standing waves; and the related concept of linearity, connected to the superposition of tones. Next, we derive the partial differential equation which governs the motion of the vibrating string. Finally, we will use what we learned about the physical nature of the problem (expressed mathematically) to solve the equation. In the last section, we use the same approach to study the problem of heat diffusion.

Given the introductory nature of this chapter and the subject matter covered, our presentation cannot be based on purely mathematical reasoning. Rather, it proceeds by plausibility arguments and aims to provide the motivation for the further rigorous analysis in the succeeding chapters. The impatient reader who wishes to begin immediately with the theorems of the subject may prefer to pass directly to the next chapter.

1 The vibrating string

The problem consists of the study of the motion of a string fixed at its end points and allowed to vibrate freely. We have in mind physical systems such as the strings of a musical instrument. As we mentioned above, we begin with a brief description of several observable physical phenomena on which our study is based. These are:

- simple harmonic motion,
- standing and traveling waves,
- harmonics and superposition of tones.

Understanding the empirical facts behind these phenomena will motivate our mathematical approach to vibrating strings.

Simple harmonic motion

Simple harmonic motion describes the behavior of the most basic oscillatory system (called the simple harmonic oscillator), and is therefore a natural place to start the study of vibrations. Consider a mass $\{m\}$ attached to a horizontal spring, which itself is attached to a fixed wall, and assume that the system lies on a frictionless surface.

Choose an axis whose origin coincides with the center of the mass when it is at rest (that is, the spring is neither stretched nor compressed), as shown in Figure 1. When the mass is displaced from its initial equilibrium

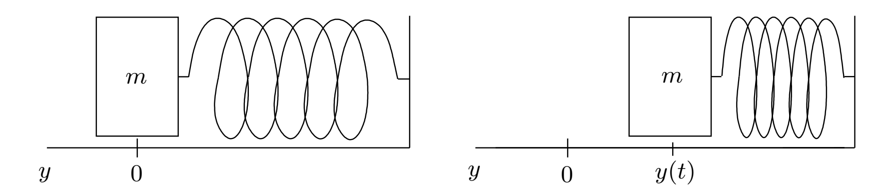


Figure 1. Simple harmonic oscillator

position and then released, it will undergo **simple harmonic motion**. This motion can be described mathematically once we have found the differential equation that governs the movement of the mass.

Let $y(t)$ denote the displacement of the mass at time t. We assume that the spring is ideal, in the sense that it satisfies Hooke's law: the restoring force F exerted by the spring on the mass is given by $F = -ky(t)$. Here

$k > 0$ is a given physical quantity called the spring constant. Applying Newton's law (force = mass × acceleration), we obtain

$$-ky(t) = my''(t),$$

where we use the notation y'' to denote the second derivative of y with respect to t. With $c = \sqrt{k/m}$, this second order ordinary differential equation becomes

$$y''(t) + c^2 y(t) = 0. \tag{1}$$

The general solution of equation (1) is given by

$$y(t) = a \cos ct + b \sin ct\,,$$

where a and b are constants. Clearly, all functions of this form solve equation (1), and Exercise 6 outlines a proof that these are the only (twice differentiable) solutions of that differential equation.

In the above expression for $y(t)$, the quantity c is given, but a and b can be any real numbers. In order to determine the particular solution of the equation, we must impose two initial conditions in view of the two unknown constants a and b. For example, if we are given $y(0)$ and $y'(0)$, the initial position and velocity of the mass, then the solution of the physical problem is unique and given by

$$y(t) = y(0) \cos ct + \frac{y'(0)}{c} \sin ct\,.$$

One can easily verify that there exist constants $A > 0$ and $\varphi \in \mathbb{R}$ such that

$$a \cos ct + b \sin ct = A \cos(ct - \varphi).$$

Because of the physical interpretation given above, one calls $A = \sqrt{a^2 + b^2}$ the "amplitude" of the motion, c its "natural frequency," φ its "phase" (uniquely determined up to an integer multiple of 2π), and $2\pi/c$ the "period" of the motion.

The typical graph of the function $A\cos(ct - \varphi)$, illustrated in Figure 2, exhibits a wavelike pattern that is obtained from translating and stretching (or shrinking) the usual graph of $\cos t$.

We make two observations regarding our examination of simple harmonic motion. The first is that the mathematical description of the most elementary oscillatory system, namely simple harmonic motion, involves

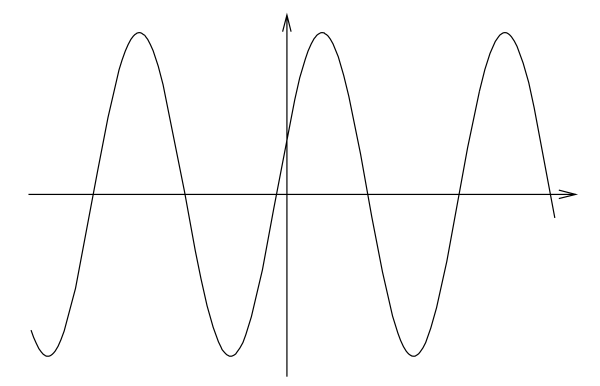


Figure 2. The graph of $A\cos(ct - \varphi)$

the most basic trigonometric functions $\cos t$ and $\sin t$. It will be important in what follows to recall the connection between these functions and complex numbers, as given in Euler's identity $e^{it} = \cos t + i \sin t$. The second observation is that simple harmonic motion is determined as a function of time by two initial conditions, one determining the position, and the other the velocity (specified, for example, at time $t = 0$). This property is shared by more general oscillatory systems, as we shall see below.

Standing and traveling waves

As it turns out, the vibrating string can be viewed in terms of one-dimensional wave motions. Here we want to describe two kinds of motions that lend themselves to simple graphic representations.

- First, we consider **standing waves**. These are wavelike motions described by the graphs $y = u(x, t)$ developing in time t as shown in Figure 3.

  In other words, there is an initial profile $y = \varphi(x)$ representing the wave at time $t = 0$, and an amplifying factor $\psi(t)$, depending on t, so that $y = u(x, t)$ with

  $$u(x, t) = \varphi(x)\psi(t).$$

  The nature of standing waves suggests the mathematical idea of "separation of variables," to which we will return later.

- A second type of wave motion that is often observed in nature is that of a **traveling wave**. Its description is particularly simple:



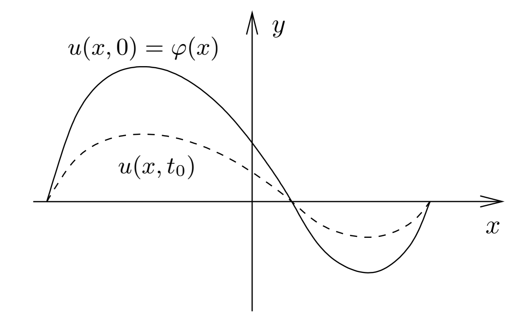


Figure 3. A standing wave at different moments in time: $t = 0$ and $t = t_0$

there is an initial profile $F(x)$ so that $u(x,t)$ equals $F(x)$ when $t = 0$. As t evolves, this profile is displaced to the right by ct units, where c is a positive constant, namely

$$u(x,t) = F(x - ct).$$

Graphically, the situation is depicted in Figure 4.

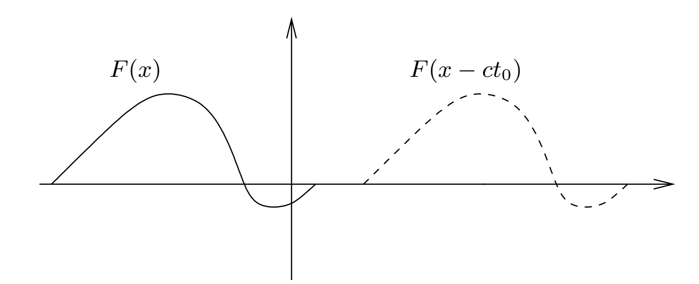


Figure 4. A traveling wave at two different moments in time: $t = 0$ and $t = t_0$

Since the movement in t is at the rate c, that constant represents the **velocity** of the wave. The function $F(x - ct)$ is a one-dimensional traveling wave moving to the right. Similarly, $u(x,t) = F(x + ct)$ is a one-dimensional traveling wave moving to the left.

Harmonics and superposition of tones

The final physical observation we want to mention (without going into any details now) is one that musicians have been aware of since time immemorial. It is the existence of harmonics, or overtones. The **pure tones** are accompanied by combinations of **overtones** which are primarily responsible for the timbre (or tone color) of the instrument. The idea of combination or superposition of tones is implemented mathematically by the basic concept of linearity, as we shall see below.

We now turn our attention to our main problem, that of describing the motion of a vibrating string. First, we derive the wave equation, that is, the partial differential equation that governs the motion of the string.

1.1 Derivation of the wave equation

Imagine a homogeneous string placed in the (x, y)-plane, and stretched along the x-axis between $x = 0$ and $x = L$. If it is set to vibrate, its displacement $y = u(x, t)$ is then a function of x and t, and the goal is to derive the differential equation which governs this function.

For this purpose, we consider the string as being subdivided into a large number N of masses (which we think of as individual particles) distributed uniformly along the x-axis, so that the nth particle has its x-coordinate at $x_n = nL/N$. We shall therefore conceive of the vibrating string as a complex system of N particles, each oscillating in the *vertical direction only*; however, unlike the simple harmonic oscillator we considered previously, each particle will have its oscillation linked to its immediate neighbor by the tension of the string.

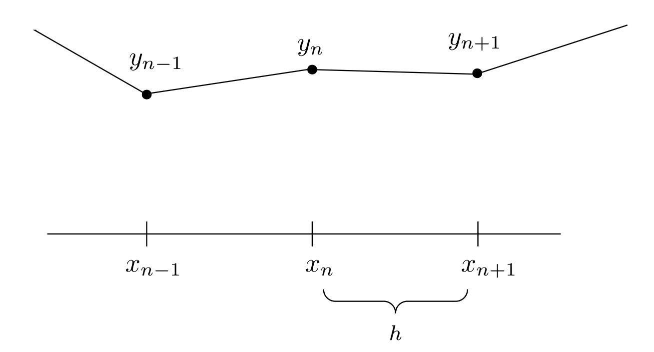


Figure 5. A vibrating string as a discrete system of masses

We then set $y_n(t) = u(x_n, t)$, and note that $x_{n+1} - x_n = h$, with $h = L/N$. If we assume that the string has constant density $\rho > 0$, it is reasonable to assign mass equal to ρh to each particle. By Newton's law, $\rho h y_n''(t)$ equals the force acting on the nth particle. We now make the simple assumption that this force is due to the effect of the two nearby particles, the ones with x-coordinates at x_{n-1} and x_{n+1} (see Figure 5). We further assume that the force (or tension) coming from the right of the nth particle is proportional to $(y_{n+1} - y_n)/h$, where h is the distance between x_{n+1} and x_n; hence we can write the tension as

$$\left(\frac{\tau}{h}\right)(y_{n+1} - y_n),$$

where $\tau > 0$ is a constant equal to the coefficient of tension of the string. There is a similar force coming from the left, and it is

$$\left(\frac{\tau}{h}\right)(y_{n-1} - y_n).$$

Altogether, adding these forces gives us the desired relation between the oscillators $y_n(t)$, namely

$$\rho h y_n''(t) = \frac{\tau}{h}\left\{y_{n+1}(t) + y_{n-1}(t) - 2y_n(t)\right\}. \tag{2}$$

On the one hand, with the notation chosen above, we see that

$$y_{n+1}(t) + y_{n-1}(t) - 2y_n(t) = u(x_n + h, t) + u(x_n - h, t) - 2u(x_n, t).$$

On the other hand, for any reasonable function $F(x)$ (that is, one that has continuous second derivatives) we have

$$\frac{F(x+h) + F(x-h) - 2F(x)}{h^2} \to F''(x) \quad \text{as } h \to 0.$$

Thus we may conclude, after dividing by h in (2) and letting h tend to zero (that is, N goes to infinity), that

$$\rho \frac{\partial^2 u}{\partial t^2} = \tau \frac{\partial^2 u}{\partial x^2},$$

or

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad \text{with } c = \sqrt{\tau/\rho}.$$

This relation is known as the **one-dimensional wave equation**, or more simply as the **wave equation**. For reasons that will be apparent later, the coefficient $c > 0$ is called the **velocity** of the motion.

In connection with this partial differential equation, we make an important simplifying mathematical remark. This has to do with **scaling**, or in the language of physics, a "change of units." That is, we can think of the coordinate x as $x = aX$ where a is an appropriate positive constant. Now, in terms of the new coordinate X, the interval $0 \leq x \leq L$ becomes $0 \leq X \leq L/a$. Similarly, we can replace the time coordinate t by $t = bT$, where b is another positive constant. If we set $U(X,T) = u(x,t)$, then

$$\frac{\partial U}{\partial X} = a\frac{\partial u}{\partial x}, \qquad \frac{\partial^2 U}{\partial X^2} = a^2\frac{\partial^2 u}{\partial x^2},$$

and similarly for the derivatives in t. So if we choose a and b appropriately, we can transform the one-dimensional wave equation into

$$\frac{\partial^2 U}{\partial T^2} = \frac{\partial^2 U}{\partial X^2},$$

which has the effect of setting the velocity c equal to 1. Moreover, we have the freedom to transform the interval $0 \leq x \leq L$ to $0 \leq X \leq \pi$. (We shall see that the choice of π is convenient in many circumstances.) All this is accomplished by taking $a = L/\pi$ and $b = L/(c\pi)$. Once we solve the new equation, we can of course return to the original equation by making the inverse change of variables. Hence, we do not sacrifice generality by thinking of the wave equation as given on the interval $[0, \pi]$ with velocity $c = 1$.

1.2 Solution to the wave equation

Having derived the equation for the vibrating string, we now explain two methods to solve it:

- using traveling waves,
- using the superposition of standing waves.

While the first approach is very simple and elegant, it does not directly give full insight into the problem; the second method accomplishes that, and moreover is of wide applicability. It was first believed that the second method applied only in the simple cases where the initial position and velocity of the string were themselves given as a superposition of standing waves. However, as a consequence of Fourier's ideas, it became clear that the problem could be worked either way for all initial conditions.

Traveling waves

To simplify matters as before, we assume that $c = 1$ and $L = \pi$, so that the equation we wish to solve becomes

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \quad \text{on } 0 \leq x \leq \pi.$$

The crucial observation is the following: if F is any twice differentiable function, then $u(x,t) = F(x+t)$ and $u(x,t) = F(x-t)$ solve the wave equation. The verification of this is a simple exercise in differentiation. Note that the graph of $u(x,t) = F(x-t)$ at time $t = 0$ is simply the graph of F, and that at time $t = 1$ it becomes the graph of F translated to the right by 1. Therefore, we recognize that $F(x-t)$ is a traveling wave which travels to the right with speed 1. Similarly, $u(x,t) = F(x+t)$ is a wave traveling to the left with speed 1. These motions are depicted in Figure 6.

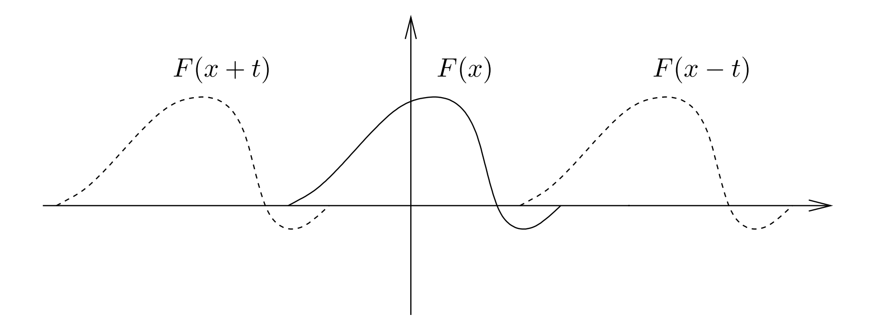


Figure 6. Waves traveling in both directions

Our discussion of tones and their combinations leads us to observe that the wave equation is **linear**. This means that if $u(x,t)$ and $v(x,t)$ are particular solutions, then so is $\alpha u(x,t) + \beta v(x,t)$, where α and β are any constants. Therefore, we may superpose two waves traveling in opposite directions to find that whenever F and G are twice differentiable functions, then

$$u(x,t) = F(x+t) + G(x-t)$$

is a solution of the wave equation. In fact, we now show that all solutions take this form.

We drop for the moment the assumption that $0 \leq x \leq \pi$, and suppose that u is a twice differentiable function which solves the wave equation

for all real x and t. Consider the following new set of variables $\xi = x + t$, $\eta = x - t$, and define $v(\xi, \eta) = u(x, t)$. The change of variables formula shows that v satisfies

$$\frac{\partial^2 v}{\partial \xi \partial \eta} = 0.$$

Integrating this relation twice gives $v(\xi, \eta) = F(\xi) + G(\eta)$, which then implies

$$u(x, t) = F(x + t) + G(x - t),$$

for some functions F and G.

We must now connect this result with our original problem, that is, the physical motion of a string. There, we imposed the restrictions $0 \leq x \leq \pi$, the initial shape of the string $u(x, 0) = f(x)$, and also the fact that the string has fixed end points, namely $u(0, t) = u(\pi, t) = 0$ for all t. To use the simple observation above, we first extend f to all of $\mathbb{R}$ by making it odd[1] on $[-\pi, \pi]$, and then periodic[2] in x of period 2π, and similarly for $u(x, t)$, the solution of our problem. Then the extension u solves the wave equation on all of $\mathbb{R}$, and $u(x, 0) = f(x)$ for all $x \in \mathbb{R}$. Therefore, $u(x, t) = F(x + t) + G(x - t)$, and setting $t = 0$ we find that

$$F(x) + G(x) = f(x).$$

Since many choices of F and G will satisfy this identity, this suggests imposing another initial condition on u (similar to the two initial conditions in the case of simple harmonic motion), namely the initial velocity of the string which we denote by $g(x)$:

$$\frac{\partial u}{\partial t}(x, 0) = g(x),$$

where of course $g(0) = g(\pi) = 0$. Again, we extend g to $\mathbb{R}$ first by making it odd over $[-\pi, \pi]$, and then periodic of period 2π. The two initial conditions of position and velocity now translate into the following system:

$$\begin{cases} F(x) + G(x) = f(x)\,, \\ F'(x) - G'(x) = g(x)\,. \end{cases}$$

[1] A function f defined on a set U is **odd** if $-x \in U$ whenever $x \in U$ and $f(-x) = -f(x)$, and **even** if $f(-x) = f(x)$.

[2] A function f on $\mathbb{R}$ is **periodic** of period ω if $f(x + \omega) = f(x)$ for all x.

Differentiating the first equation and adding it to the second, we obtain

$$2F'(x) = f'(x) + g(x).$$

Similarly

$$2G'(x) = f'(x) - g(x),$$

and hence there are constants C_1 and C_2 so that

$$F(x) = \frac{1}{2}\left[f(x) + \int_0^x g(y)\,dy\right] + C_1$$

and

$$G(x) = \frac{1}{2}\left[f(x) - \int_0^x g(y)\,dy\right] + C_2.$$

Since $F(x) + G(x) = f(x)$ we conclude that $C_1 + C_2 = 0$, and therefore, our final solution of the wave equation with the given initial conditions takes the form

$$u(x,t) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{x-t}^{x+t} g(y)\,dy.$$

The form of this solution is known as **d'Alembert's formula**. Observe that the extensions we chose for f and g guarantee that the string always has fixed ends, that is, $u(0,t) = u(\pi,t) = 0$ for all t.

A final remark is in order. The passage from $t \geq 0$ to $t \in \mathbb{R}$, and then back to $t \geq 0$, which was made above, exhibits the time reversal property of the wave equation. In other words, a solution u to the wave equation for $t \geq 0$, leads to a solution u^- defined for negative time $t < 0$ simply by setting $u^-(x,t) = u(x,-t)$, a fact which follows from the invariance of the wave equation under the transformation $t \mapsto -t$. The situation is quite different in the case of the heat equation.

Superposition of standing waves

We turn to the second method of solving the wave equation, which is based on two fundamental conclusions from our previous physical observations. By our considerations of standing waves, we are led to look for special solutions to the wave equation which are of the form $\varphi(x)\psi(t)$. This procedure, which works equally well in other contexts (in the case of the heat equation, for instance), is called **separation of variables** and constructs solutions that are called pure tones. Then by the linearity

of the wave equation, we can expect to combine these pure tones into a more complex combination of sound. Pushing this idea further, we can hope ultimately to express the general solution of the wave equation in terms of sums of these particular solutions.

Note that one side of the wave equation involves only differentiation in x, while the other, only differentiation in t. This observation provides another reason to look for solutions of the equation in the form $u(x,t) = \varphi(x)\psi(t)$ (that is, to "separate variables"), the hope being to reduce a difficult partial differential equation into a system of simpler ordinary differential equations. In the case of the wave equation, with u of the above form, we get

$$\varphi(x)\psi''(t) = \varphi''(x)\psi(t),$$

and therefore

$$\frac{\psi''(t)}{\psi(t)} = \frac{\varphi''(x)}{\varphi(x)}.$$

The key observation here is that the left-hand side depends only on t, and the right-hand side only on x. This can happen only if both sides are equal to a constant, say λ. Therefore, the wave equation reduces to the following

$$\left\{ \begin{array}{l} \psi''(t) - \lambda\psi(t) \ = \ 0 \\ \varphi''(x) - \lambda\varphi(x) = \ 0. \end{array} \right. \tag{3}$$

We focus our attention on the first equation in the above system. At this point, the reader will recognize the equation we obtained in the study of simple harmonic motion. Note that we need to consider only the case when $\lambda < 0$, since when $\lambda \geq 0$ the solution ψ will not oscillate as time varies. Therefore, we may write $\lambda = -m^2$, and the solution of the equation is then given by

$$\psi(t) = A\cos mt + B\sin mt.$$

Similarly, we find that the solution of the second equation in (3) is

$$\varphi(x) = \tilde{A}\cos mx + \tilde{B}\sin mx.$$

Now we take into account that the string is attached at $x = 0$ and $x = \pi$. This translates into $\varphi(0) = \varphi(\pi) = 0$, which in turn gives $\tilde{A} = 0$, and if $\tilde{B} \neq 0$, then m must be an integer. If $m = 0$, the solution vanishes identically, and if $m \leq -1$, we may rename the constants and reduce to

the case $m \geq 1$ since the function $\sin y$ is odd and $\cos y$ is even. Finally, we arrive at the guess that for each $m \geq 1$, the function

$$u_m(x,t) = (A_m \cos mt + B_m \sin mt) \sin mx,$$

which we recognize as a **standing wave**, is a solution to the wave equation. Note that in the above argument we divided by φ and ψ, which sometimes vanish, so one must actually check by hand that the standing wave u_m solves the equation. This straightforward calculation is left as an exercise to the reader.

Before proceeding further with the analysis of the wave equation, we pause to discuss standing waves in more detail. The terminology comes from looking at the graph of $u_m(x,t)$ for each fixed t. Suppose first that $m = 1$, and take $u(x,t) = \cos t \sin x$. Then, Figure 7 (a) gives the graph of u for different values of t.

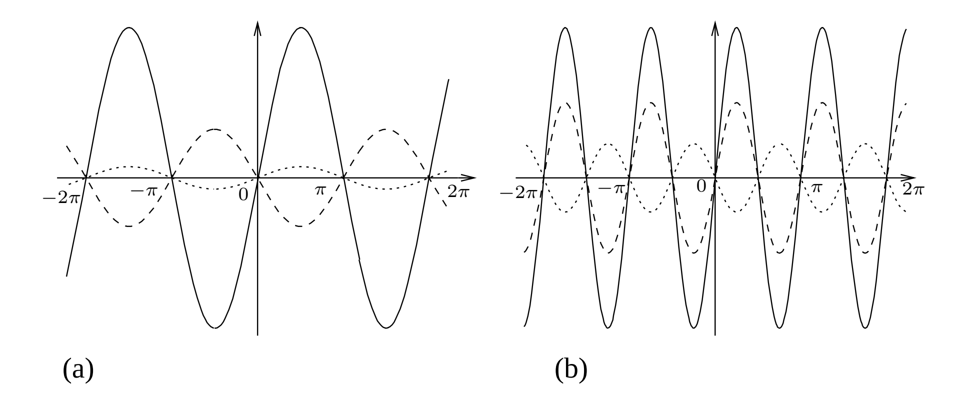


Figure 7. Fundamental tone (a) and overtones (b) at different moments in time

The case $m = 1$ corresponds to the **fundamental tone** or **first harmonic** of the vibrating string.

We now take $m = 2$ and look at $u(x,t) = \cos 2t \sin 2x$. This corresponds to the **first overtone** or **second harmonic**, and this motion is described in Figure 7 (b). Note that $u(\pi/2, t) = 0$ for all t. Such points, which remain motionless in time, are called nodes, while points whose motion has maximum amplitude are named anti-nodes.

For higher values of m we get more overtones or higher harmonics. Note that as m increases, the frequency increases, and the period $2\pi/m$

decreases. Therefore, the fundamental tone has a lower frequency than the overtones.

We now return to the original problem. Recall that the wave equation is linear in the sense that if u and v solve the equation, so does $\alpha u + \beta v$ for any constants α and β. This allows us to construct more solutions by taking linear combinations of the standing waves u_m. This technique, called **superposition**, leads to our final guess for a solution of the wave equation

$$u(x,t) = \sum_{m=1}^{\infty} (A_m \cos mt + B_m \sin mt) \sin mx. \tag{4}$$

Note that the above sum is infinite, so that questions of convergence arise, but since most of our arguments so far are formal, we will not worry about this point now.

Suppose the above expression gave *all* the solutions to the wave equation. If we then require that the initial position of the string at time $t = 0$ is given by the shape of the graph of the function f on $[0, \pi]$, with of course $f(0) = f(\pi) = 0$, we would have $u(x, 0) = f(x)$, hence

$$\sum_{m=1}^{\infty} A_m \sin mx = f(x).$$

Since the initial shape of the string can be any reasonable function f, we must ask the following basic question:

> Given a function f on $[0, \pi]$ (with $f(0) = f(\pi) = 0$), can we find coefficients A_m so that

$$f(x) = \sum_{m=1}^{\infty} A_m \sin mx \ ? \tag{5}$$

This question is stated loosely, but a lot of our effort in the next two chapters of this book will be to formulate the question precisely and attempt to answer it. This was the basic problem that initiated the study of Fourier analysis.

A simple observation allows us to guess a formula giving A_m if the expansion (5) were to hold. Indeed, we multiply both sides by $\sin nx$

and integrate between $[0, \pi]$; working formally, we obtain

$$\int_0^\pi f(x) \sin nx \, dx = \int_0^\pi \left(\sum_{m=1}^{\infty} A_m \sin mx \right) \sin nx \, dx$$
$$= \sum_{m=1}^{\infty} A_m \int_0^\pi \sin mx \, \sin nx \, dx = A_n \cdot \frac{\pi}{2},$$

where we have used the fact that

$$\int_0^\pi \sin mx \, \sin nx \, dx = \begin{cases} 0 & \text{if } m \neq n, \\ \pi/2 & \text{if } m = n. \end{cases}$$

Therefore, the guess for A_n, called the nth Fourier sine coefficient of f, is

$$A_n = \frac{2}{\pi} \int_0^\pi f(x) \sin nx \, dx. \tag{6}$$

We shall return to this formula, and other similar ones, later.

One can transform the question about Fourier sine series on $[0, \pi]$ to a more general question on the interval $[-\pi, \pi]$. If we could express f on $[0, \pi]$ in terms of a sine series, then this expansion would also hold on $[-\pi, \pi]$ if we extend f to this interval by making it odd. Similarly, one can ask if an even function $g(x)$ on $[-\pi, \pi]$ can be expressed as a cosine series, namely

$$g(x) = \sum_{m=0}^{\infty} A'_m \cos mx.$$

More generally, since an arbitrary function F on $[-\pi, \pi]$ can be expressed as $f + g$, where f is odd and g is even,[3] we may ask if F can be written as

$$F(x) = \sum_{m=1}^{\infty} A_m \sin mx + \sum_{m=0}^{\infty} A'_m \cos mx,$$

or by applying Euler's identity $e^{ix} = \cos x + i \sin x$, we could hope that F takes the form

$$F(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx}.$$

[3]Take, for example, $f(x) = [F(x) - F(-x)]/2$ and $g(x) = [F(x) + F(-x)]/2$.

By analogy with (6), we can use the fact that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}e^{-inx}\,dx = \begin{cases} 0 & \text{if } n \neq m \\ 1 & \text{if } n = m, \end{cases}$$

to see that one expects that

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)e^{-inx}\,dx.$$

The quantity a_n is called the nth **Fourier coefficient** of F.

We can now reformulate the problem raised above:

> **Question:** Given any reasonable function F on $[-\pi, \pi]$, with Fourier coefficients defined above, is it true that

$$F(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx} \ ? \tag{7}$$

This formulation of the problem, in terms of complex exponentials, is the form we shall use the most in what follows.

Joseph Fourier (1768-1830) was the first to believe that an "arbitrary" function F could be given as a series (7). In other words, his idea was that any function is the linear combination (possibly infinite) of the most basic trigonometric functions $\sin mx$ and $\cos mx$, where m ranges over the integers.[4] Although this idea was implicit in earlier work, Fourier had the conviction that his predecessors lacked, and he used it in his study of heat diffusion; this began the subject of "Fourier analysis." This discipline, which was first developed to solve certain physical problems, has proved to have many applications in mathematics and other fields as well, as we shall see later.

We return to the wave equation. To formulate the problem correctly, we must impose two initial conditions, as our experience with simple harmonic motion and traveling waves indicated. The conditions assign the initial position and velocity of the string. That is, we require that u satisfy the differential equation and the two conditions

$$u(x,0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x,0) = g(x),$$

[4]The first proof that a general class of functions can be represented by Fourier series was given later by Dirichlet; see Problem 6, Chapter 4.

where f and g are pre-assigned functions. Note that this is consistent with (4) in that this requires that f and g be expressible as

$$f(x) = \sum_{m=1}^{\infty} A_m \sin mx \quad \text{and} \quad g(x) = \sum_{m=1}^{\infty} mB_m \sin mx.$$

1.3 Example: the plucked string

We now apply our reasoning to the particular problem of the plucked string. For simplicity we choose units so that the string is taken on the interval $[0, \pi]$, and it satisfies the wave equation with $c = 1$. The string is assumed to be plucked to height h at the point p with $0 < p < \pi$; this is the initial position. That is, we take as our initial position the triangular shape given by

$$f(x) = \begin{cases} \dfrac{xh}{p} & \text{for } 0 \le x \le p \\[2ex] \dfrac{h(\pi - x)}{\pi - p} & \text{for } p \le x \le \pi, \end{cases}$$

which is depicted in Figure 8.

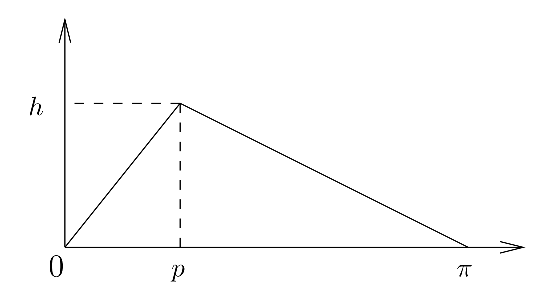


Figure 8. Initial position of a plucked string

We also choose an initial velocity $g(x)$ identically equal to 0. Then, we can compute the Fourier coefficients of f (Exercise 9), and assuming that the answer to the question raised before (5) is positive, we obtain

$$f(x) = \sum_{m=1}^{\infty} A_m \sin mx \quad \text{with} \quad A_m = \frac{2h}{m^2} \frac{\sin mp}{p(\pi - p)}.$$

Thus

$$u(x,t) = \sum_{m=1}^{\infty} A_m \cos mt \, \sin mx, \tag{8}$$

and note that this series converges absolutely. The solution can also be expressed in terms of traveling waves. In fact

$$u(x,t) = \frac{f(x+t) + f(x-t)}{2}. \tag{9}$$

Here $f(x)$ is defined for all x as follows: first, f is extended to $[-\pi, \pi]$ by making it odd, and then f is extended to the whole real line by making it periodic of period 2π, that is, $f(x + 2\pi k) = f(x)$ for all integers k.

Observe that (8) implies (9) in view of the trigonometric identity

$$\cos v \sin u = \frac{1}{2}\,[\sin(u+v) + \sin(u-v)].$$

As a final remark, we should note an unsatisfactory aspect of the solution to this problem, which however is in the nature of things. Since the initial data $f(x)$ for the plucked string is not twice continuously differentiable, neither is the function u (given by (9)). Hence u is not truly a solution of the wave equation: while $u(x,t)$ does represent the position of the plucked string, it does not satisfy the partial differential equation we set out to solve! This state of affairs may be understood properly only if we realize that u does solve the equation, but in an appropriate generalized sense. A better understanding of this phenomenon requires ideas relevant to the study of "weak solutions" and the theory of "distributions." These topics we consider only later, in Books III and IV.

2 The heat equation

We now discuss the problem of heat diffusion by following the same framework as for the wave equation. First, we derive the time-dependent heat equation, and then study the steady-state heat equation in the disc, which leads us back to the basic question (7).

2.1 Derivation of the heat equation

Consider an infinite metal plate which we model as the plane $\mathbb{R}^2$, and suppose we are given an initial heat distribution at time $t = 0$. Let the temperature at the point (x, y) at time t be denoted by $u(x, y, t)$.

Consider a small square centered at (x_0, y_0) with sides parallel to the axis and of side length h, as shown in Figure 9. The amount of heat energy in S at time t is given by

$$H(t) = \sigma \iint_S u(x, y, t)\, dx\, dy\,,$$

where $\sigma > 0$ is a constant called the specific heat of the material. Therefore, the heat flow into S is

$$\frac{\partial H}{\partial t} = \sigma \iint_S \frac{\partial u}{\partial t}\, dx\, dy\,,$$

which is approximately equal to

$$\sigma h^2 \frac{\partial u}{\partial t}(x_0, y_0, t),$$

since the area of S is h^2. Now we apply Newton's law of cooling, which states that heat flows from the higher to lower temperature at a rate proportional to the difference, that is, the gradient.

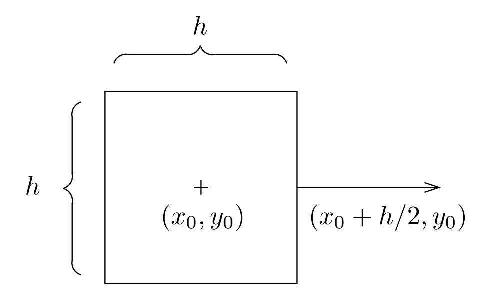


Figure 9. Heat flow through a small square

The heat flow through the vertical side on the right is therefore

$$-\kappa h \frac{\partial u}{\partial x}(x_0 + h/2, y_0, t)\,,$$

where $\kappa > 0$ is the conductivity of the material. A similar argument for the other sides shows that the total heat flow through the square S is

given by

$$\kappa h \left[\frac{\partial u}{\partial x}(x_0 + h/2, y_0, t) - \frac{\partial u}{\partial x}(x_0 - h/2, y_0, t) \right.$$
$$\left. + \frac{\partial u}{\partial y}(x_0, y_0 + h/2, t) - \frac{\partial u}{\partial y}(x_0, y_0 - h/2, t) \right].$$

Applying the mean value theorem and letting h tend to zero, we find that

$$\frac{\sigma}{\kappa}\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2};$$

this is called the **time-dependent heat equation**, often abbreviated to the heat equation.

2.2 Steady-state heat equation in the disc

After a long period of time, there is no more heat exchange, so that the system reaches thermal equilibrium and $\partial u/\partial t = 0$. In this case, the time-dependent heat equation reduces to the **steady-state heat equation**

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \tag{10}$$

The operator $\partial^2/\partial x^2 + \partial^2/\partial y^2$ is of such importance in mathematics and physics that it is often abbreviated as $\triangle$ and given a name: the Laplace operator or **Laplacian**. So the steady-state heat equation is written as

$$\triangle u = 0,$$

and solutions to this equation are called **harmonic functions**.

Consider the unit disc in the plane

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\},$$

whose boundary is the unit circle C. In polar coordinates (r, θ), with $0 \le r$ and $0 \le \theta < 2\pi$, we have

$$D = \{(r, \theta) : 0 \le r < 1\} \quad \text{and} \quad C = \{(r, \theta) : r = 1\}.$$

The problem, often called the **Dirichlet problem** (for the Laplacian on the unit disc), is to solve the steady-state heat equation in the unit

disc subject to the boundary condition $u = f$ on C. This corresponds to fixing a predetermined temperature distribution on the circle, waiting a long time, and then looking at the temperature distribution inside the disc.

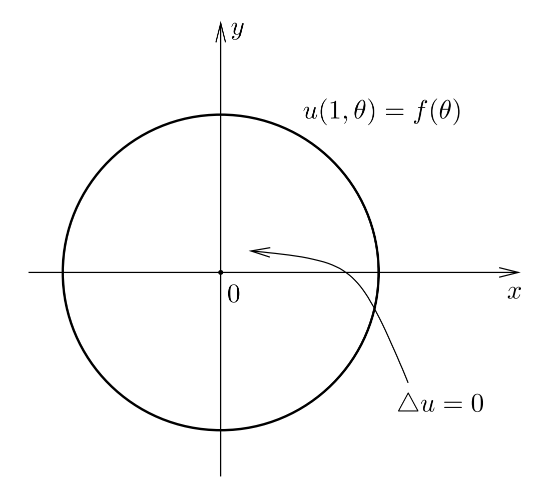


Figure 10. The Dirichlet problem for the disc

While the method of separation of variables will turn out to be useful for equation (10), a difficulty comes from the fact that the boundary condition is not easily expressed in terms of rectangular coordinates. Since this boundary condition is best described by the coordinates (r, θ), namely $u(1, \theta) = f(\theta)$, we rewrite the Laplacian in polar coordinates. An application of the chain rule gives (Exercise 10):

$$\triangle u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}.$$

We now multiply both sides by r^2, and since $\triangle u = 0$, we get

$$r^2 \frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} = -\frac{\partial^2 u}{\partial \theta^2}.$$

Separating these variables, and looking for a solution of the form $u(r, \theta) = F(r)G(\theta)$, we find

$$\frac{r^2 F''(r) + rF'(r)}{F(r)} = -\frac{G''(\theta)}{G(\theta)}.$$

Since the two sides depend on different variables, they must both be constant, say equal to λ. We therefore get the following equations:

$$\begin{cases} G''(\theta) + \lambda G(\theta) = 0\,, \\ r^2 F''(r) + rF'(r) - \lambda F(r) = 0. \end{cases}$$

Since G must be periodic of period 2π, this implies that $\lambda \geq 0$ and (as we have seen before) that $\lambda = m^2$ where m is an integer; hence

$$G(\theta) = \tilde{A}\cos m\theta + \tilde{B}\sin m\theta.$$

An application of Euler's identity, $e^{ix} = \cos x + i\sin x$, allows one to rewrite G in terms of complex exponentials,

$$G(\theta) = Ae^{im\theta} + Be^{-im\theta}.$$

With $\lambda = m^2$ and $m \neq 0$, two simple solutions of the equation in F are $F(r) = r^m$ and $F(r) = r^{-m}$ (Exercise 11 gives further information about these solutions). If $m = 0$, then $F(r) = 1$ and $F(r) = \log r$ are two solutions. If $m > 0$, we note that r^{-m} grows unboundedly large as r tends to zero, so $F(r)G(\theta)$ is unbounded at the origin; the same occurs when $m = 0$ and $F(r) = \log r$. We reject these solutions as contrary to our intuition. Therefore, we are left with the following special functions:

$$u_m(r,\theta) = r^{|m|}e^{im\theta}, \qquad m \in \mathbb{Z}.$$

We now make the important observation that (10) is *linear*, and so as in the case of the vibrating string, we may superpose the above special solutions to obtain the presumed general solution:

$$u(r,\theta) = \sum_{m=-\infty}^{\infty} a_m r^{|m|} e^{im\theta}.$$

If this expression gave all the solutions to the steady-state heat equation, then for a reasonable f we should have

$$u(1,\theta) = \sum_{m=-\infty}^{\infty} a_m e^{im\theta} = f(\theta).$$

We therefore ask again in this context: given any reasonable function f on $[0, 2\pi]$ with $f(0) = f(2\pi)$, can we find coefficients a_m so that

$$f(\theta) = \sum_{m=-\infty}^{\infty} a_m e^{im\theta} \quad ?$$

Historical Note: D'Alembert (in 1747) first solved the equation of the vibrating string using the method of traveling waves. This solution was elaborated by Euler a year later. In 1753, D. Bernoulli proposed the solution which for all intents and purposes is the Fourier series given by (4), but Euler was not entirely convinced of its full generality, since this could hold only if an "arbitrary" function could be expanded in Fourier series. D'Alembert and other mathematicians also had doubts. This viewpoint was changed by Fourier (in 1807) in his study of the heat equation, where his conviction and work eventually led others to a complete proof that a general function could be represented as a Fourier series.

3 Exercises

1. If $z = x + iy$ is a complex number with $x, y \in \mathbb{R}$, we define

$$|z| = (x^2 + y^2)^{1/2}$$

and call this quantity the **modulus** or **absolute value** of z.

(a) What is the geometric interpretation of $|z|$?

(b) Show that if $|z| = 0$, then $z = 0$.

(c) Show that if $\lambda \in \mathbb{R}$, then $|\lambda z| = |\lambda||z|$, where $|\lambda|$ denotes the standard absolute value of a real number.

(d) If z_1 and z_2 are two complex numbers, prove that

$$|z_1 z_2| = |z_1||z_2| \quad \text{and} \quad |z_1 + z_2| \leq |z_1| + |z_2|.$$

(e) Show that if $z \neq 0$, then $|1/z| = 1/|z|$.

2. If $z = x + iy$ is a complex number with $x, y \in \mathbb{R}$, we define the **complex conjugate** of z by

$$\overline{z} = x - iy.$$

(a) What is the geometric interpretation of $\overline{z}$?

(b) Show that $|z|^2 = z\overline{z}$.

(c) Prove that if z belongs to the unit circle, then $1/z = \overline{z}$.

3. A sequence of complex numbers $\{w_n\}_{n=1}^{\infty}$ is said to converge if there exists $w \in \mathbb{C}$ such that

$$\lim_{n\to\infty} |w_n - w| = 0,$$

and we say that w is a limit of the sequence.

(a) Show that a converging sequence of complex numbers has a unique limit.

The sequence $\{w_n\}_{n=1}^{\infty}$ is said to be a **Cauchy sequence** if for every $\epsilon > 0$ there exists a positive integer N such that

$$|w_n - w_m| < \epsilon \qquad \text{whenever } n, m > N.$$

(b) Prove that a sequence of complex numbers converges if and only if it is a Cauchy sequence. [Hint: A similar theorem exists for the convergence of a sequence of real numbers. Why does it carry over to sequences of complex numbers?]

A series $\sum_{n=1}^{\infty} z_n$ of complex numbers is said to converge if the sequence formed by the partial sums

$$S_N = \sum_{n=1}^{N} z_n$$

converges. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of non-negative real numbers such that the series $\sum_n a_n$ converges.

(c) Show that if $\{z_n\}_{n=1}^{\infty}$ is a sequence of complex numbers satisfying $|z_n| \leq a_n$ for all n, then the series $\sum_n z_n$ converges. [Hint: Use the Cauchy criterion.]

4. For $z \in \mathbb{C}$, we define the **complex exponential** by

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

(a) Prove that the above definition makes sense, by showing that the series converges for every complex number z. Moreover, show that the convergence is uniform[5] on every bounded subset of $\mathbb{C}$.

(b) If z_1, z_2 are two complex numbers, prove that $e^{z_1}e^{z_2} = e^{z_1+z_2}$. [Hint: Use the binomial theorem to expand $(z_1 + z_2)^n$, as well as the formula for the binomial coefficients.]

[5]A sequence of functions $\{f_n(z)\}_{n=1}^{\infty}$ is said to be uniformly convergent on a set S if there exists a function f on S so that for every $\epsilon > 0$ there is an integer N such that $|f_n(z) - f(z)| < \epsilon$ whenever $n > N$ and $z \in S$.

(c) Show that if z is purely imaginary, that is, $z = iy$ with $y \in \mathbb{R}$, then

$$e^{iy} = \cos y + i \sin y.$$

This is Euler's identity. [Hint: Use power series.]

(d) More generally,

$$e^{x+iy} = e^{x}(\cos y + i \sin y)$$

whenever $x, y \in \mathbb{R}$, and show that

$$|e^{x+iy}| = e^{x}.$$

(e) Prove that $e^{z} = 1$ if and only if $z = 2\pi k i$ for some integer k.

(f) Show that every complex number $z = x + iy$ can be written in the form

$$z = re^{i\theta},$$

where r is unique and in the range $0 \leq r < \infty$, and $\theta \in \mathbb{R}$ is unique up to an integer multiple of 2π. Check that

$$r = |z| \quad \text{and} \quad \theta = \arctan(y/x)$$

whenever these formulas make sense.

(g) In particular, $i = e^{i\pi/2}$. What is the geometric meaning of multiplying a complex number by i? Or by $e^{i\theta}$ for any $\theta \in \mathbb{R}$?

(h) Given $\theta \in \mathbb{R}$, show that

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

These are also called Euler's identities.

(i) Use the complex exponential to derive trigonometric identities such as

$$\cos(\theta + \vartheta) = \cos\theta\cos\vartheta - \sin\theta\sin\vartheta,$$

and then show that

$$\begin{aligned} 2\sin\theta\sin\varphi &= \cos(\theta - \varphi) - \cos(\theta + \varphi), \\ 2\sin\theta\cos\varphi &= \sin(\theta + \varphi) + \sin(\theta - \varphi). \end{aligned}$$

This calculation connects the solution given by d'Alembert in terms of traveling waves and the solution in terms of superposition of standing waves.

5. Verify that $f(x) = e^{inx}$ is periodic with period 2π and that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\,dx = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \neq 0. \end{cases}$$

Use this fact to prove that if $n, m \geq 1$ we have

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos nx \, \cos mx \, dx = \begin{cases} 0 & \text{if } n \neq m, \\ 1 & n = m, \end{cases}$$

and similarly

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \sin nx \, \sin mx \, dx = \begin{cases} 0 & \text{if } n \neq m, \\ 1 & n = m. \end{cases}$$

Finally, show that

$$\int_{-\pi}^{\pi} \sin nx \, \cos mx \, dx = 0 \qquad \text{for any } n, m.$$

[Hint: Calculate $e^{inx}e^{-imx} + e^{inx}e^{imx}$ and $e^{inx}e^{-imx} - e^{inx}e^{imx}$.]

6. Prove that if f is a twice continuously differentiable function on $\mathbb{R}$ which is a solution of the equation

$$f''(t) + c^2 f(t) = 0,$$

then there exist constants a and b such that

$$f(t) = a\cos ct + b\sin ct.$$

This can be done by differentiating the two functions $g(t) = f(t)\cos ct - c^{-1}f'(t)\sin ct$ and $h(t) = f(t)\sin ct + c^{-1}f'(t)\cos ct$.

7. Show that if a and b are real, then one can write

$$a\cos ct + b\sin ct = A\cos(ct - \varphi),$$

where $A = \sqrt{a^2 + b^2}$, and φ is chosen so that

$$\cos\varphi = \frac{a}{\sqrt{a^2+b^2}} \quad \text{and} \quad \sin\varphi = \frac{b}{\sqrt{a^2+b^2}}.$$

8. Suppose F is a function on (a, b) with two continuous derivatives. Show that whenever x and $x + h$ belong to (a, b), one may write

$$F(x + h) = F(x) + hF'(x) + \frac{h^2}{2}F''(x) + h^2\varphi(h)\,,$$

where $\varphi(h) \to 0$ as $h \to 0$.

Deduce that

$$\frac{F(x+h)+F(x-h)-2F(x)}{h^2} \to F''(x) \qquad \text{as } h \to 0.$$

[Hint: This is simply a Taylor expansion. It may be obtained by noting that

$$F(x+h) - F(x) = \int_x^{x+h} F'(y)\,dy,$$

and then writing $F'(y) = F'(x) + (y-x)F''(x) + (y-x)\psi(y-x)$, where $\psi(h) \to 0$ as $h \to 0$.]

9. In the case of the plucked string, use the formula for the Fourier sine coefficients to show that

$$A_m = \frac{2h}{m^2} \frac{\sin mp}{p(\pi - p)}.$$

For what position of p are the second, fourth, ... harmonics missing? For what position of p are the third, sixth, ... harmonics missing?

10. Show that the expression of the Laplacian

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

is given in polar coordinates by the formula

$$\triangle = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}.$$

Also, prove that

$$\left|\frac{\partial u}{\partial x}\right|^2 + \left|\frac{\partial u}{\partial y}\right|^2 = \left|\frac{\partial u}{\partial r}\right|^2 + \frac{1}{r^2}\left|\frac{\partial u}{\partial \theta}\right|^2.$$

11. Show that if $n \in \mathbb{Z}$ the only solutions of the differential equation

$$r^2 F''(r) + rF'(r) - n^2 F(r) = 0,$$

which are twice differentiable when $r > 0$, are given by linear combinations of r^n and r^{-n} when $n \neq 0$, and 1 and $\log r$ when $n = 0$.

[Hint: If F solves the equation, write $F(r) = g(r)r^n$, find the equation satisfied by g, and conclude that $rg'(r) + 2ng(r) = c$ where c is a constant.]

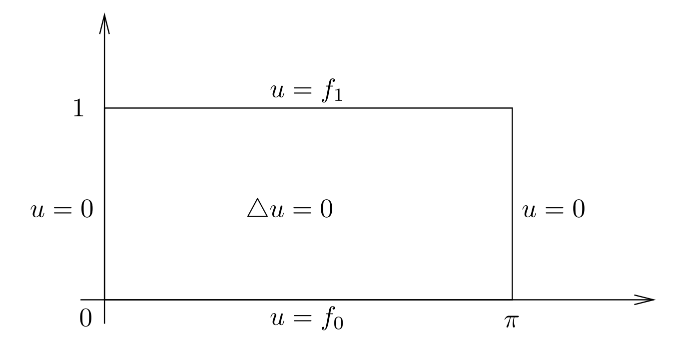


Figure 11. Dirichlet problem in a rectangle

4 Problem

1. Consider the Dirichlet problem illustrated in Figure 11.

More precisely, we look for a solution of the steady-state heat equation $\triangle u = 0$ in the rectangle $R = \{(x, y) : \ 0 \leq x \leq \pi, \ \ 0 \leq y \leq 1\}$ that vanishes on the vertical sides of R, and so that

$$u(x, 0) = f_0(x) \quad \text{ and } \quad u(x, 1) = f_1(x)\,,$$

where f_0 and f_1 are initial data which fix the temperature distribution on the horizontal sides of the rectangle.

Use separation of variables to show that if f_0 and f_1 have Fourier expansions

$$f_0(x) = \sum_{k=1}^{\infty} A_k \sin kx \quad \text{ and } \quad f_1(x) = \sum_{k=1}^{\infty} B_k \sin kx,$$

then

$$u(x, y) = \sum_{k=1}^{\infty} \left(\frac{\sinh k(1-y)}{\sinh k} A_k + \frac{\sinh ky}{\sinh k} B_k \right) \sin kx.$$

We recall the definitions of the hyperbolic sine and cosine functions:

$$\sinh x = \frac{e^x - e^{-x}}{2} \quad \text{ and } \quad \cosh x = \frac{e^x + e^{-x}}{2}.$$

Compare this result with the solution of the Dirichlet problem in the strip obtained in Problem 3, Chapter 5.

2 Basic Properties of Fourier Series

> Nearly fifty years had passed without any progress on the question of analytic representation of an arbitrary function, when an assertion of Fourier threw new light on the subject. Thus a new era began for the development of this part of Mathematics and this was heralded in a stunning way by major developments in mathematical Physics.
>
> *B. Riemann,* 1854

In this chapter, we begin our rigorous study of Fourier series. We set the stage by introducing the main objects in the subject, and then formulate some basic problems which we have already touched upon earlier.

Our first result disposes of the question of uniqueness: Are two functions with the same Fourier coefficients necessarily equal? Indeed, a simple argument shows that if both functions are continuous, then in fact they must agree.

Next, we take a closer look at the partial sums of a Fourier series. Using the formula for the Fourier coefficients (which involves an integration), we make the key observation that these sums can be written conveniently as integrals:

$$\frac{1}{2\pi}\int D_N(x-y)f(y)\,dy,$$

where $\{D_N\}$ is a family of functions called the Dirichlet kernels. The above expression is the convolution of f with the function D_N. Convolutions will play a critical role in our analysis. In general, given a family of functions $\{K_n\}$, we are led to investigate the limiting properties as n tends to infinity of the convolutions

$$\frac{1}{2\pi}\int K_n(x-y)f(y)\,dy.$$

We find that if the family $\{K_n\}$ satisfies the three important properties of "good kernels," then the convolutions above tend to $f(x)$ as $n \to \infty$ (at least when f is continuous). In this sense, the family $\{K_n\}$ is an

"approximation to the identity." Unfortunately, the Dirichlet kernels D_N do not belong to the category of good kernels, which indicates that the question of convergence of Fourier series is subtle.

Instead of pursuing at this stage the problem of convergence, we consider various other methods of summing the Fourier series of a function. The first method, which involves averages of partial sums, leads to convolutions with good kernels, and yields an important theorem of Fejér. From this, we deduce the fact that a continuous function on the circle can be approximated uniformly by trigonometric polynomials. Second, we may also sum the Fourier series in the sense of Abel and again encounter a family of good kernels. In this case, the results about convolutions and good kernels lead to a solution of the Dirichlet problem for the steady-state heat equation in the disc, considered at the end of the previous chapter.

1 Examples and formulation of the problem

We commence with a brief description of the types of functions with which we shall be concerned. Since the Fourier coefficients of f are defined by

$$a_n = \frac{1}{L}\int_0^L f(x)e^{-2\pi inx/L}\,dx, \qquad \text{for } n \in \mathbb{Z},$$

where f is complex-valued on $[0, L]$, it will be necessary to place some integrability conditions on f. We shall therefore assume for the remainder of this book that all functions are at least Riemann integrable.[1] Sometimes it will be illuminating to focus our attention on functions that are more "regular," that is, functions that possess certain continuity or differentiability properties. Below, we list several classes of functions in increasing order of generality. We emphasize that we will not generally restrict our attention to real-valued functions, contrary to what the following pictures may suggest; we will almost always allow functions that take values in the complex numbers $\mathbb{C}$. Furthermore, we sometimes think of our functions as being defined on the circle rather than an interval. We elaborate upon this below.

[1] Limiting ourselves to Riemann integrable functions is natural at this elementary stage of study of the subject. The more advanced notion of Lebesgue integrability will be taken up in Book III.

Everywhere continuous functions

These are the complex-valued functions f which are continuous at every point of the segment $[0, L]$. A typical continuous function is sketched in Figure 1 (a). We shall note later that continuous functions on the circle satisfy the additional condition $f(0) = f(L)$.

Piecewise continuous functions

These are bounded functions on $[0, L]$ which have only finitely many discontinuities. An example of such a function with simple discontinuities is pictured in Figure 1 (b).

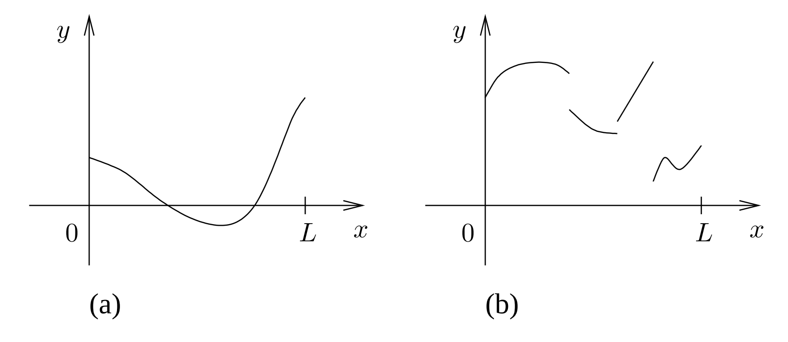


Figure 1. Functions on $[0, L]$: continuous and piecewise continuous

This class of functions is wide enough to illustrate many of the theorems in the next few chapters. However, for logical completeness we consider also the more general class of Riemann integrable functions. This more extended setting is natural since the formula for the Fourier coefficients involves integration.

Riemann integrable functions

This is the most general class of functions we will be concerned with. Such functions are bounded, but may have infinitely many discontinuities. We recall the definition of integrability. A real-valued function f defined on $[0, L]$ is **Riemann integrable** (which we abbreviate as **integrable**[2]) if it is *bounded*, and if for every $\epsilon > 0$, there is a subdivision $0 = x_0 < x_1 < \cdots < x_{N-1} < x_N = L$ of the interval $[0, L]$, so that if $\mathcal{U}$

[2]Starting in Book III, the term "integrable" will be used in the broader sense of Lebesgue theory.

and $\mathcal{L}$ are, respectively, the upper and lower sums of f for this subdivision, namely

$$\mathcal{U} = \sum_{j=1}^{N} [\sup_{x_{j-1} \leq x \leq x_j} f(x)](x_j - x_{j-1})$$

and

$$\mathcal{L} = \sum_{j=1}^{N} [\inf_{x_{j-1} \leq x \leq x_j} f(x)](x_j - x_{j-1}),$$

then we have $\mathcal{U} - \mathcal{L} < \epsilon$. Finally, we say that a complex-valued function is integrable if its real and imaginary parts are integrable. It is worthwhile to remember at this point that the sum and product of two integrable functions are integrable.

A simple example of an integrable function on $[0, 1]$ with infinitely many discontinuities is given by

$$f(x) = \begin{cases} 1 & \text{if } 1/(n+1) < x \leq 1/n \text{ and } n \text{ is odd,} \\ 0 & \text{if } 1/(n+1) < x \leq 1/n \text{ and } n \text{ is even,} \\ 0 & \text{if } x = 0. \end{cases}$$

This example is illustrated in Figure 2. Note that f is discontinuous when $x = 1/n$ and at $x = 0$.

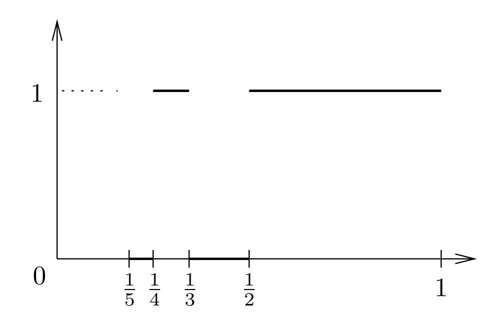


Figure 2. A Riemann integrable function

More elaborate examples of integrable functions whose discontinuities are dense in the interval $[0, 1]$ are described in Problem 1. In general, while integrable functions may have infinitely many discontinuities, these

functions are actually characterized by the fact that, in a precise sense, their discontinuities are not too numerous: they are "negligible," that is, the set of points where an integrable function is discontinuous has "measure 0." The reader will find further details about Riemann integration in the appendix.

From now on, we shall always assume that our functions are integrable, even if we do not state this requirement explicitly.

Functions on the circle

There is a natural connection between 2π-periodic functions on $\mathbb{R}$ like the exponentials $e^{in\theta}$, functions on an interval of length 2π, and functions on the unit circle. This connection arises as follows.

A point on the unit circle takes the form $e^{i\theta}$, where θ is a real number that is unique up to integer multiples of 2π. If F is a function on the circle, then we may define for each real number θ

$$f(\theta) = F(e^{i\theta}),$$

and observe that with this definition, the function f is periodic on $\mathbb{R}$ of period 2π, that is, $f(\theta + 2\pi) = f(\theta)$ for all θ. The integrability, continuity and other smoothness properties of F are determined by those of f. For instance, we say that F is integrable on the circle if f is integrable on every interval of length 2π. Also, F is continuous on the circle if f is continuous on $\mathbb{R}$, which is the same as saying that f is continuous on any interval of length 2π. Moreover, F is continuously differentiable if f has a continuous derivative, and so forth.

Since f has period 2π, we may restrict it to any interval of length 2π, say $[0, 2\pi]$ or $[-\pi, \pi]$, and still capture the initial function F on the circle. We note that f must take the same value at the end-points of the interval since they correspond to the same point on the circle. Conversely, any function on $[0, 2\pi]$ for which $f(0) = f(2\pi)$ can be extended to a periodic function on $\mathbb{R}$ which can then be identified as a function on the circle. In particular, a continuous function f on the interval $[0, 2\pi]$ gives rise to a continuous function on the circle if and only if $f(0) = f(2\pi)$.

In conclusion, functions on $\mathbb{R}$ that 2π-periodic, and functions on an interval of length 2π that take on the same value at its end-points, are two equivalent descriptions of the same mathematical objects, namely, functions on the circle.

In this connection, we mention an item of notational usage. When our functions are defined on an interval on the line, we often use x as the independent variable; however, when we consider these as functions

on the circle, we usually replace the variable x by θ. As the reader will note, we are not strictly bound by this rule since this practice is mostly a matter of convenience.

1.1 Main definitions and some examples

We now begin our study of Fourier analysis with the precise definition of the Fourier series of a function. Here, it is important to pin down where our function is originally defined. If f is an integrable function given on an interval $[a, b]$ of length L (that is, $b - a = L$), then the nth **Fourier coefficient** of f is defined by

$$\hat{f}(n) = \frac{1}{L}\int_a^b f(x)e^{-2\pi inx/L}\,dx, \qquad n \in \mathbb{Z}.$$

The **Fourier series** of f is given formally[3] by

$$\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx/L}.$$

We shall sometimes write a_n for the Fourier coefficients of f, and use the notation

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{2\pi inx/L}$$

to indicate that the series on the right-hand side is the Fourier series of f.

For instance, if f is an integrable function on the interval $[-\pi, \pi]$, then the nth Fourier coefficient of f is

$$\hat{f}(n) = a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}\,d\theta, \qquad n \in \mathbb{Z},$$

and the Fourier series of f is

$$f(\theta) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

Here we use θ as a variable since we think of it as an angle ranging from $-\pi$ to π.

[3] At this point, we do not say anything about the convergence of the series.

Also, if f is defined on $[0, 2\pi]$, then the formulas are the same as above, except that we integrate from 0 to 2π in the definition of the Fourier coefficients.

We may also consider the Fourier coefficients and Fourier series for a function defined on the circle. By our previous discussion, we may think of a function on the circle as a function f on $\mathbb{R}$ which is 2π-periodic. We may restrict the function f to any interval of length 2π, for instance $[0, 2\pi]$ or $[-\pi, \pi]$, and compute its Fourier coefficients. Fortunately, f is *periodic* and Exercise 1 shows that the resulting integrals are independent of the chosen interval. Thus the Fourier coefficients of a function on the circle are well defined.

Finally, we shall sometimes consider a function g given on $[0, 1]$. Then

$$\hat{g}(n) = a_n = \int_0^1 g(x)e^{-2\pi inx}\, dx \quad \text{and} \quad g(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{2\pi inx}.$$

Here we use x for a variable ranging from 0 to 1.

Of course, if f is initially given on $[0, 2\pi]$, then $g(x) = f(2\pi x)$ is defined on $[0, 1]$ and a change of variables shows that the nth Fourier coefficient of f equals the nth Fourier coefficient of g.

Fourier series are part of a larger family called the **trigonometric series** which, by definition, are expressions of the form $\sum_{n=-\infty}^{\infty} c_n e^{2\pi inx/L}$ where $c_n \in \mathbb{C}$. If a trigonometric series involves only finitely many non-zero terms, that is, $c_n = 0$ for all large $|n|$, it is called a **trigonometric polynomial**; its **degree** is the largest value of $|n|$ for which $c_n \neq 0$.

The Nth **partial sum** of the Fourier series of f, for N a positive integer, is a particular example of a trigonometric polynomial. It is given by

$$S_N(f)(x) = \sum_{n=-N}^{N} \hat{f}(n)e^{2\pi inx/L}.$$

Note that by definition, the above sum is *symmetric* since n ranges from $-N$ to N, a choice that is natural because of the resulting decomposition of the Fourier series as sine and cosine series. As a consequence, the convergence of Fourier series will be understood (in this book) as the "limit" as N tends to infinity of these symmetric sums.

In fact, using the partial sums of the Fourier series, we can reformulate the basic question raised in Chapter 1 as follows:

Problem: In what sense does $S_N(f)$ converge to f as $N \to \infty$?

Before proceeding further with this question, we turn to some simple examples of Fourier series.

EXAMPLE 1. Let $f(\theta) = \theta$ for $-\pi \leq \theta \leq \pi$. The calculation of the Fourier coefficients requires a simple integration by parts. First, if $n \neq 0$, then

$$\begin{aligned}\hat{f}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \theta e^{-in\theta}\, d\theta \\ &= \frac{1}{2\pi}\left[-\frac{\theta}{in}e^{-in\theta}\right]_{-\pi}^{\pi} + \frac{1}{2\pi in}\int_{-\pi}^{\pi} e^{-in\theta}\, d\theta \\ &= \frac{(-1)^{n+1}}{in},\end{aligned}$$

and if $n = 0$ we clearly have

$$\hat{f}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \theta\, d\theta = 0.$$

Hence, the Fourier series of f is given by

$$f(\theta) \sim \sum_{n\neq 0} \frac{(-1)^{n+1}}{in} e^{in\theta} = 2\sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin n\theta}{n}.$$

The first sum is over all non-zero integers, and the second is obtained by an application of Euler's identities. It is possible to prove by elementary means that the above series converges for every θ, but it is not obvious that it converges to $f(\theta)$. This will be proved later (Exercises 8 and 9 deal with a similar situation).

EXAMPLE 2. Define $f(\theta) = (\pi - \theta)^2/4$ for $0 \leq \theta \leq 2\pi$. Then successive integration by parts similar to that performed in the previous example yield

$$f(\theta) \sim \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{\cos n\theta}{n^2}.$$

EXAMPLE 3. The Fourier series of the function

$$f(\theta) = \frac{\pi}{\sin \pi\alpha} e^{i(\pi-\theta)\alpha}$$

on $[0, 2\pi]$ is

$$f(\theta) \sim \sum_{n=-\infty}^{\infty} \frac{e^{in\theta}}{n+\alpha},$$

whenever α is not an integer.

EXAMPLE 4. The trigonometric polynomial defined for $x \in [-\pi, \pi]$ by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}$$

is called the Nth **Dirichlet kernel** and is of fundamental importance in the theory (as we shall see later). Notice that its Fourier coefficients a_n have the property that $a_n = 1$ if $|n| \leq N$ and $a_n = 0$ otherwise. A closed form formula for the Dirichlet kernel is

$$D_N(x) = \frac{\sin((N + \frac{1}{2})x)}{\sin(x/2)}.$$

This can be seen by summing the geometric progressions

$$\sum_{n=0}^{N} \omega^n \quad \text{and} \quad \sum_{n=-N}^{-1} \omega^n$$

with $\omega = e^{ix}$. These sums are, respectively, equal to

$$\frac{1 - \omega^{N+1}}{1 - \omega} \quad \text{and} \quad \frac{\omega^{-N} - 1}{1 - \omega}.$$

Their sum is then

$$\frac{\omega^{-N} - \omega^{N+1}}{1 - \omega} = \frac{\omega^{-N-1/2} - \omega^{N+1/2}}{\omega^{-1/2} - \omega^{1/2}} = \frac{\sin((N + \frac{1}{2})x)}{\sin(x/2)},$$

giving the desired result.

EXAMPLE 5. The function $P_r(\theta)$, called the **Poisson kernel**, is defined for $\theta \in [-\pi, \pi]$ and $0 \leq r < 1$ by the absolutely and uniformly convergent series

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}.$$

This function arose implicitly in the solution of the steady-state heat equation on the unit disc discussed in Chapter 1. Note that in calculating the Fourier coefficients of $P_r(\theta)$ we can interchange the order of integration and summation since the sum converges uniformly in θ for

each fixed r, and obtain that the nth Fourier coefficient equals $r^{|n|}$. One can also sum the series for $P_r(\theta)$ and see that

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}.$$

In fact,

$$P_r(\theta) = \sum_{n=0}^{\infty} \omega^n + \sum_{n=1}^{\infty} \overline{\omega}^n \qquad \text{with } \omega = re^{i\theta},$$

where both series converge absolutely. The first sum (an infinite geometric progression) equals $1/(1 - \omega)$, and likewise, the second is $\overline{\omega}/(1 - \overline{\omega})$. Together, they combine to give

$$\frac{1 - \overline{\omega} + (1 - \omega)\overline{\omega}}{(1 - \omega)(1 - \overline{\omega})} = \frac{1 - |\omega|^2}{|1 - \omega|^2} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2},$$

as claimed. The Poisson kernel will reappear later in the context of Abel summability of the Fourier series of a function.

Let us return to the problem formulated earlier. The definition of the Fourier series of f is purely formal, and it is not obvious whether it converges to f. In fact, the solution of this problem can be very hard, or relatively easy, depending on the sense in which we expect the series to converge, or on what additional restrictions we place on f.

Let us be more precise. Suppose, for the sake of this discussion, that the function f (which is always assumed to be Riemann integrable) is defined on $[-\pi, \pi]$. The first question one might ask is whether the partial sums of the Fourier series of f converge to f pointwise. That is, do we have

$$\text{(1)} \qquad \lim_{N\to\infty} S_N(f)(\theta) = f(\theta) \qquad \text{for every } \theta?$$

We see quite easily that in general we cannot expect this result to be true at every θ, since we can always change an integrable function at one point without changing its Fourier coefficients. As a result, we might ask the same question assuming that f is continuous and periodic. For a long time it was believed that under these additional assumptions the answer would be "yes." It was a surprise when Du Bois-Reymond showed that there exists a continuous function whose Fourier series diverges at a point. We will give such an example in the next chapter. Despite this negative result, we might ask what happens if we add more smoothness conditions on f: for example, we might assume that f is continuously

differentiable, or twice continuously differentiable. We will see that then the Fourier series of f converges to f uniformly.

We will also interpret the limit (1) by showing that the Fourier series sums, in the sense of Cesàro or Abel, to the function f at all of its points of continuity. This approach involves appropriate averages of the partial sums of the Fourier series of f.

Finally, we can also define the limit (1) in the mean square sense. In the next chapter, we will show that if f is merely integrable, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|S_N(f)(\theta)-f(\theta)|^2\,d\theta \to 0 \quad \text{as } N\to\infty.$$

It is of interest to know that the problem of pointwise convergence of Fourier series was settled in 1966 by L. Carleson, who showed, among other things, that if f is integrable in our sense,[4] then the Fourier series of f converges to f except possibly on a set of "measure 0." The proof of this theorem is difficult and beyond the scope of this book.

2 Uniqueness of Fourier series

If we were to assume that the Fourier series of functions f converge to f in an appropriate sense, then we could infer that a function is uniquely determined by its Fourier coefficients. This would lead to the following statement: if f and g have the same Fourier coefficients, then f and g are necessarily equal. By taking the difference $f-g$, this proposition can be reformulated as: if $\hat{f}(n)=0$ for all $n\in\mathbb{Z}$, then $f=0$. As stated, this assertion cannot be correct without reservation, since calculating Fourier coefficients requires integration, and we see that, for example, any two functions which differ at finitely many points have the same Fourier series. However, we do have the following positive result.

Theorem 2.1 *Suppose that f is an integrable function on the circle with $\hat{f}(n)=0$ for all $n\in\mathbb{Z}$. Then $f(\theta_0)=0$ whenever f is continuous at the point θ_0.*

Thus, in terms of what we know about the set of discontinuities of integrable functions,[5] we can conclude that f vanishes for "most" values of θ.

Proof. We suppose first that f is real-valued, and argue by contradiction. Assume, without loss of generality, that f is defined on

---

[4] Carleson's proof actually holds for the wider class of functions which are square integrable in the Lebesgue sense.

[5] See the appendix.

$[-\pi, \pi]$, that $\theta_0 = 0$, and $f(0) > 0$. The idea now is to construct a family of trigonometric polynomials $\{p_k\}$ that "peak" at 0, and so that $\int p_k(\theta) f(\theta)\, d\theta \to \infty$ as $k \to \infty$. This will be our desired contradiction since these integrals are equal to zero by assumption.

Since f is continuous at 0, we can choose $0 < \delta \leq \pi/2$, so that $f(\theta) > f(0)/2$ whenever $|\theta| < \delta$. Let

$$p(\theta) = \epsilon + \cos\theta,$$

where $\epsilon > 0$ is chosen so small that $|p(\theta)| < 1 - \epsilon/2$, whenever $\delta \leq |\theta| \leq \pi$. Then, choose a positive η with $\eta < \delta$, so that $p(\theta) \geq 1 + \epsilon/2$, for $|\theta| < \eta$. Finally, let

$$p_k(\theta) = [p(\theta)]^k,$$

and select B so that $|f(\theta)| \leq B$ for all θ. This is possible since f is integrable, hence bounded. Figure 3 illustrates the family $\{p_k\}$. By

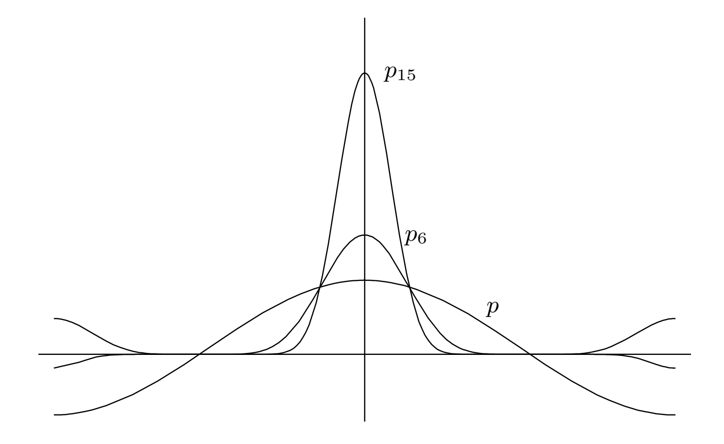


Figure 3. The functions p, p_6, and p_{15} when $\epsilon = 0.1$

construction, each p_k is a trigonometric polynomial, and since $\hat{f}(n) = 0$ for all n, we must have

$$\int_{-\pi}^{\pi} f(\theta) p_k(\theta)\, d\theta = 0 \qquad \text{for all } k.$$

However, we have the estimate

$$\left| \int_{\delta \leq |\theta|} f(\theta) p_k(\theta)\, d\theta \right| \leq 2\pi B (1 - \epsilon/2)^k.$$

Also, our choice of δ guarantees that $p(\theta)$ and $f(\theta)$ are non-negative whenever $|\theta| < \delta$, thus

$$\int_{\eta \leq |\theta| < \delta} f(\theta) p_k(\theta)\, d\theta \geq 0.$$

Finally,

$$\int_{|\theta| < \eta} f(\theta) p_k(\theta)\, d\theta \geq 2\eta \frac{f(0)}{2} (1 + \epsilon/2)^k.$$

Therefore, $\int p_k(\theta) f(\theta)\, d\theta \to \infty$ as $k \to \infty$, and this concludes the proof when f is real-valued. In general, write $f(\theta) = u(\theta) + iv(\theta)$, where u and v are real-valued. If we define $\overline{f}(\theta) = \overline{f(\theta)}$, then

$$u(\theta) = \frac{f(\theta) + \overline{f}(\theta)}{2} \quad \text{and} \quad v(\theta) = \frac{f(\theta) - \overline{f}(\theta)}{2i},$$

and since $\hat{\overline{f}}(n) = \overline{\hat{f}(-n)}$, we conclude that the Fourier coefficients of u and v all vanish, hence $f = 0$ at its points of continuity. The idea

of constructing a family of functions (trigonometric polynomials in this case) which peak at the origin, together with other nice properties, will play an important role in this book. Such families of functions will be taken up later in Section 4 in connection with the notion of convolution. For now, note that the above theorem implies the following.

Corollary 2.2 *If f is continuous on the circle and $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$.*

The next corollary shows that the problem (1) formulated earlier has a simple positive answer under the assumption that the series of Fourier coefficients converges absolutely.

Corollary 2.3 *Suppose that f is a continuous function on the circle and that the Fourier series of f is absolutely convergent, $\sum_{n=-\infty}^{\infty} |\hat{f}(n)| < \infty$. Then, the Fourier series converges uniformly to f, that is,*

$$\lim_{N \to \infty} S_N(f)(\theta) = f(\theta) \quad \text{uniformly in } \theta.$$

Proof. Recall that if a sequence of continuous functions converges uniformly, then the limit is also continuous. Now observe that the assumption $\sum |\hat{f}(n)| < \infty$ implies that the partial sums of the Fourier

series of f converge absolutely and uniformly, and therefore the function g defined by

$$g(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{in\theta} = \lim_{N\to\infty} \sum_{n=-N}^{N} \hat{f}(n) e^{in\theta}$$

is continuous on the circle. Moreover, the Fourier coefficients of g are precisely $\hat{f}(n)$ since we can interchange the infinite sum with the integral (a consequence of the uniform convergence of the series). Therefore, the previous corollary applied to the function $f - g$ yields $f = g$, as desired.

What conditions on f would guarantee the absolute convergence of its Fourier series? As it turns out, the smoothness of f is directly related to the decay of the Fourier coefficients, and in general, the smoother the function, the faster this decay. As a result, we can expect that relatively smooth functions equal their Fourier series. This is in fact the case, as we now show.

In order to state the result concisely we introduce the standard ***"O"* notation**, which we will use freely in the rest of this book. For example, the statement $\hat{f}(n) = O(1/|n|^2)$ as $|n| \to \infty$, means that the left-hand side is bounded by a constant multiple of the right-hand side; that is, there exists $C > 0$ with $|\hat{f}(n)| \leq C/|n|^2$ for all large $|n|$. More generally, $f(x) = O(g(x))$ as $x \to a$ means that for some constant C, $|f(x)| \leq C|g(x)|$ as x approaches a. In particular, $f(x) = O(1)$ means that f is bounded.

Corollary 2.4 *Suppose that f is a twice continuously differentiable function on the circle. Then*

$$\hat{f}(n) = O(1/|n|^2) \quad \text{as } |n| \to \infty,$$

so that the Fourier series of f converges absolutely and uniformly to f.

Proof. The estimate on the Fourier coefficients is proved by integrating by parts twice for $n \neq 0$. We obtain

$$\begin{aligned} 2\pi \hat{f}(n) &= \int_0^{2\pi} f(\theta) e^{-in\theta}\, d\theta \\ &= \left[f(\theta) \cdot \frac{-e^{-in\theta}}{in} \right]_0^{2\pi} + \frac{1}{in} \int_0^{2\pi} f'(\theta) e^{-in\theta}\, d\theta \\ &= \frac{1}{in} \int_0^{2\pi} f'(\theta) e^{-in\theta}\, d\theta \\ &= \frac{1}{in} \left[f'(\theta) \cdot \frac{-e^{-in\theta}}{in} \right]_0^{2\pi} + \frac{1}{(in)^2} \int_0^{2\pi} f''(\theta) e^{-in\theta}\, d\theta \\ &= \frac{-1}{n^2} \int_0^{2\pi} f''(\theta) e^{-in\theta}\, d\theta. \end{aligned}$$

The quantities in brackets vanish since f and f' are periodic. Therefore

$$2\pi |n|^2 |\hat{f}(n)| \leq \left| \int_0^{2\pi} f''(\theta) e^{-in\theta}\, d\theta \right| \leq \int_0^{2\pi} |f''(\theta)|\, d\theta \leq C,$$

where the constant C is independent of n. (We can take $C = 2\pi B$ where B is a bound for f''.) Since $\sum 1/n^2$ converges, the proof of the corollary is complete.

Incidentally, we have also established the following important identity:

$$\widehat{f'}(n) = in \hat{f}(n), \qquad \text{for all } n \in \mathbb{Z}.$$

If $n \neq 0$ the proof is given above, and if $n = 0$ it is left as an exercise to the reader. So if f is differentiable and $f \sim \sum a_n e^{in\theta}$, then $f' \sim \sum a_n in e^{in\theta}$. Also, if f is twice continuously differentiable, then $f'' \sim \sum a_n (in)^2 e^{in\theta}$, and so on. Further smoothness conditions on f imply even better decay of the Fourier coefficients (Exercise 10).

There are also stronger versions of Corollary 2.4. It can be shown, for example, that the Fourier series of f converges absolutely, assuming only that f has one continuous derivative. Even more generally, the Fourier series of f converges absolutely (and hence uniformly to f) if f satisfies a **Hölder condition** of order α, with $\alpha > 1/2$, that is,

$$\sup_\theta |f(\theta + t) - f(\theta)| \leq A|t|^\alpha \qquad \text{for all } t.$$

For more on these matters, see the exercises at the end of Chapter 3.

At this point it is worthwhile to introduce a common notation: we say that f belongs to the **class** C^k if f is k times continuously differentiable. Belonging to the class C^k or satisfying a Hölder condition are two possible ways to describe the *smoothness* of a function.

3 Convolutions

The notion of convolution of two functions plays a fundamental role in Fourier analysis; it appears naturally in the context of Fourier series but also serves more generally in the analysis of functions in other settings.

Given two 2π-periodic integrable functions f and g on $\mathbb{R}$, we define their **convolution** $f * g$ on $[-\pi, \pi]$ by

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x-y)\, dy. \tag{2}$$

The above integral makes sense for each x, since the product of two integrable functions is again integrable. Also, since the functions are periodic, we can change variables to see that

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-y) g(y)\, dy.$$

Loosely speaking, convolutions correspond to "weighted averages." For instance, if $g = 1$ in (2), then $f * g$ is constant and equal to $\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)\, dy$, which we may interpret as the average value of f on the circle. Also, the convolution $(f * g)(x)$ plays a role similar to, and in some sense replaces, the pointwise product $f(x)g(x)$ of the two functions f and g.

In the context of this chapter, our interest in convolutions originates from the fact that the partial sums of the Fourier series of f can be expressed as follows:

$$\begin{aligned}
S_N(f)(x) &= \sum_{n=-N}^{N} \hat{f}(n) e^{inx} \\
&= \sum_{n=-N}^{N} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny}\, dy \right) e^{inx} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) \left(\sum_{n=-N}^{N} e^{in(x-y)} \right) dy \\
&= (f * D_N)(x),
\end{aligned}$$

where D_N is the Nth Dirichlet kernel (see Example 4) given by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}.$$

So we observe that the problem of understanding $S_N(f)$ reduces to the understanding of the convolution $f * D_N$.

We begin by gathering some of the main properties of convolutions.

Proposition 3.1 *Suppose that f, g, and h are 2π-periodic integrable functions. Then:*

(i) $f * (g + h) = (f * g) + (f * h)$.

(ii) $(cf) * g = c(f * g) = f * (cg)$ *for any* $c \in \mathbb{C}$.

(iii) $f * g = g * f$.

(iv) $(f * g) * h = f * (g * h)$.

(v) $f * g$ *is continuous.*

(vi) $\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n)$.

The first four points describe the algebraic properties of convolutions: linearity, commutativity, and associativity. Property (v) exhibits an important principle: the convolution of $f * g$ is "more regular" than f or g. Here, $f * g$ is continuous while f and g are merely (Riemann) integrable. Finally, (vi) is key in the study of Fourier series. In general, the Fourier coefficients of the product fg are not the product of the Fourier coefficients of f and g. However, (vi) says that this relation holds if we replace the product of the two functions f and g by their convolution $f * g$.

Proof. Properties (i) and (ii) follow at once from the linearity of the integral.

The other properties are easily deduced if we assume also that f and g are continuous. In this case, we may freely interchange the order of

integration. For instance, to establish (vi) we write

$$\begin{aligned}
\widehat{f*g}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(f*g)(x)e^{-inx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\left(\int_{-\pi}^{\pi}f(y)g(x-y)\,dy\right)e^{-inx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}f(y)e^{-iny}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}g(x-y)e^{-in(x-y)}\,dx\right)dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}f(y)e^{-iny}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}g(x)e^{-inx}\,dx\right)dy \\
&= \hat{f}(n)\hat{g}(n).
\end{aligned}$$

To prove (iii), one first notes that if F is continuous and 2π-periodic, then

$$\int_{-\pi}^{\pi}F(y)\,dy = \int_{-\pi}^{\pi}F(x-y)\,dy \qquad \text{for any } x\in\mathbb{R}.$$

The verification of this identity consists of a change of variables $y\mapsto -y$, followed by a translation $y\mapsto y-x$. Then, one takes $F(y)=f(y)g(x-y)$.

Also, (iv) follows by interchanging two integral signs, and an appropriate change of variables.

Finally, we show that if f and g are continuous, then $f*g$ is continuous. First, we may write

$$(f*g)(x_1)-(f*g)(x_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi}f(y)\left[g(x_1-y)-g(x_2-y)\right]\,dy.$$

Since g is continuous it must be uniformly continuous on any closed and bounded interval. But g is also periodic, so it must be uniformly continuous on all of $\mathbb{R}$; given $\epsilon>0$ there exists $\delta>0$ so that $|g(s)-g(t)|<\epsilon$ whenever $|s-t|<\delta$. Then, $|x_1-x_2|<\delta$ implies $|(x_1-y)-(x_2-y)|<\delta$ for any y, hence

$$\begin{aligned}
|(f*g)(x_1)-(f*g)(x_2)| &\le \frac{1}{2\pi}\left|\int_{-\pi}^{\pi}f(y)\left[g(x_1-y)-g(x_2-y)\right]\,dy\right| \\
&\le \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(y)|\,|g(x_1-y)-g(x_2-y)|\,dy \\
&\le \frac{\epsilon}{2\pi}\int_{-\pi}^{\pi}|f(y)|\,dy \\
&\le \frac{\epsilon}{2\pi}\,2\pi\,B\,,
\end{aligned}$$

where B is chosen so that $|f(x)| \leq B$ for all x. As a result, we conclude that $f * g$ is continuous, and the proposition is proved, at least when f and g are continuous.

In general, when f and g are merely integrable, we may use the results established so far (when f and g are continuous), together with the following approximation lemma, whose proof may be found in the appendix.

Lemma 3.2 *Suppose f is integrable on the circle and bounded by B. Then there exists a sequence $\{f_k\}_{k=1}^{\infty}$ of continuous functions on the circle so that*

$$\sup_{x\in[-\pi,\pi]} |f_k(x)| \leq B \quad \text{for all } k = 1, 2, \ldots,$$

and

$$\int_{-\pi}^{\pi} |f(x) - f_k(x)|\, dx \to 0 \quad \text{as } k \to \infty.$$

Using this result, we may complete the proof of the proposition as follows. Apply Lemma 3.2 to f and g to obtain sequences $\{f_k\}$ and $\{g_k\}$ of approximating continuous functions. Then

$$f * g - f_k * g_k = (f - f_k) * g + f_k * (g - g_k).$$

By the properties of the sequence $\{f_k\}$,

$$\begin{aligned} |(f - f_k) * g(x)| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x-y) - f_k(x-y)|\, |g(y)|\, dy \\ &\leq \frac{1}{2\pi} \sup_y |g(y)| \int_{-\pi}^{\pi} |f(y) - f_k(y)|\, dy \\ &\to 0 \quad \text{as } k \to \infty. \end{aligned}$$

Hence $(f - f_k) * g \to 0$ uniformly in x. Similarly, $f_k * (g - g_k) \to 0$ uniformly, and therefore $f_k * g_k$ tends uniformly to $f * g$. Since each $f_k * g_k$ is continuous, it follows that $f * g$ is also continuous, and we have (v).

Next, we establish (vi). For each fixed integer n we must have $\widehat{f_k * g_k}(n) \to \widehat{f * g}(n)$ as k tends to infinity since $f_k * g_k$ converges uniformly to $f * g$. However, we found earlier that $\widehat{f_k}(n)\widehat{g_k}(n) = \widehat{f_k * g_k}(n)$ because both f_k and g_k are continuous. Hence

$$\begin{aligned} |\hat{f}(n) - \hat{f}_k(n)| &= \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} (f(x) - f_k(x)) e^{-inx}\, dx \right| \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - f_k(x)|\, dx, \end{aligned}$$

and as a result we find that $\widehat{f_k}(n) \to \hat{f}(n)$ as k goes to infinity. Similarly $\widehat{g_k}(n) \to \hat{g}(n)$, and the desired property is established once we let k tend to infinity. Finally, properties (iii) and (iv) follow from the same kind of arguments.

4 Good kernels

In the proof of Theorem 2.1 we constructed a sequence of trigonometric polynomials $\{p_k\}$ with the property that the functions p_k peaked at the origin. As a result, we could isolate the behavior of f at the origin. In this section, we return to such families of functions, but this time in a more general setting. First, we define the notion of good kernel, and discuss the characteristic properties of such functions. Then, by the use of convolutions, we show how these kernels can be used to recover a given function.

A family of kernels $\{K_n(x)\}_{n=1}^{\infty}$ on the circle is said to be a family of **good kernels** if it satisfies the following properties:

(a) For all $n \geq 1$,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(x)\, dx = 1.$$

(b) There exists $M > 0$ such that for all $n \geq 1$,

$$\int_{-\pi}^{\pi} |K_n(x)|\, dx \leq M.$$

(c) For every $\delta > 0$,

$$\int_{\delta \leq |x| \leq \pi} |K_n(x)|\, dx \to 0, \qquad \text{as } n \to \infty.$$

In practice we shall encounter families where $K_n(x) \geq 0$, in which case (b) is a consequence of (a). We may interpret the kernels $K_n(x)$ as weight distributions on the circle: property (a) says that K_n assigns unit mass to the whole circle $[-\pi, \pi]$, and (c) that this mass concentrates near the origin as n becomes large.[6] Figure 4 (a) illustrates the typical character of a family of good kernels.

The importance of good kernels is highlighted by their use in connection with convolutions.

[6]In the limit, a family of good kernels represents the "Dirac delta function." This terminology comes from physics.



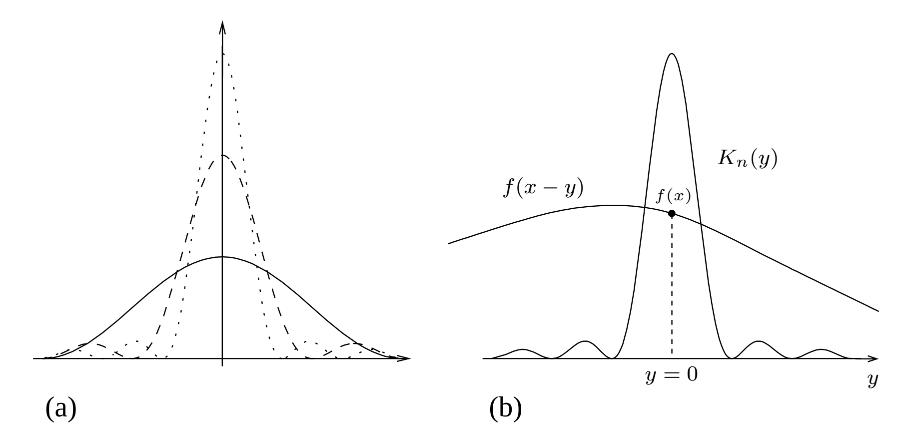


Figure 4. Good kernels

Theorem 4.1 *Let* $\{K_n\}_{n=1}^{\infty}$ *be a family of good kernels, and* f *an integrable function on the circle. Then*

$$\lim_{n\to\infty}(f * K_n)(x) = f(x)$$

whenever f *is continuous at* x*. If* f *is continuous everywhere, then the above limit is uniform.*

Because of this result, the family $\{K_n\}$ is sometimes referred to as an **approximation to the identity**.

We have previously interpreted convolutions as weighted averages. In this context, the convolution

$$(f * K_n)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)K_n(y)\,dy$$

is the average of $f(x-y)$, where the weights are given by $K_n(y)$. However, the weight distribution K_n concentrates its mass at $y = 0$ as n becomes large. Hence in the integral, the value $f(x)$ is assigned the full mass as $n \to \infty$. Figure 4 (b) illustrates this point.

Proof of Theorem 4.1. If $\epsilon > 0$ and f is continuous at x, choose δ so that $|y| < \delta$ implies $|f(x-y) - f(x)| < \epsilon$. Then, by the first property of good kernels, we can write

$$\begin{aligned}(f * K_n)(x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(y)f(x-y)\,dy - f(x)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(y)[f(x-y) - f(x)]\,dy.\end{aligned}$$

Hence,

$$\begin{aligned}|(f*K_n)(x)-f(x)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(y)[f(x-y)-f(x)]\,dy\right| \\ &\leq \frac{1}{2\pi}\int_{|y|<\delta} |K_n(y)|\,|f(x-y)-f(x)|\,dy \\ &\qquad + \frac{1}{2\pi}\int_{\delta\leq|y|\leq\pi} |K_n(y)|\,|f(x-y)-f(x)|\,dy \\ &\leq \frac{\epsilon}{2\pi}\int_{-\pi}^{\pi}|K_n(y)|\,dy + \frac{2B}{2\pi}\int_{\delta\leq|y|\leq\pi}|K_n(y)|\,dy,\end{aligned}$$

where B is a bound for f. The first term is bounded by $\epsilon M/2\pi$ because of the second property of good kernels. By the third property we see that for all large n, the second term will be less than ϵ. Therefore, for some constant $C>0$ and all large n we have

$$|(f*K_n)(x)-f(x)| \leq C\epsilon,$$

thereby proving the first assertion in the theorem. If f is continuous everywhere, then it is uniformly continuous, and δ can be chosen independent of x. This provides the desired conclusion that $f*K_n\to f$ uniformly.

Recall from the beginning of Section 3 that

$$S_N(f)(x) = (f*D_N)(x)\,,$$

where $D_N(x)=\sum_{n=-N}^{N} e^{inx}$ is the Dirichlet kernel. It is natural now for us to ask whether D_N is a good kernel, since if this were true, Theorem 4.1 would imply that the Fourier series of f converges to $f(x)$ whenever f is continuous at x. Unfortunately, this is not the case. Indeed, an estimate shows that D_N violates the second property; more precisely, one has (see Problem 2)

$$\int_{-\pi}^{\pi}|D_N(x)|\,dx \geq c\log N, \qquad \text{as } N\to\infty.$$

However, we should note that the formula for D_N as a sum of exponentials immediately gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)\,dx = 1,$$

so the first property of good kernels is actually verified. The fact that the mean value of D_N is 1, while the integral of its absolute value is large,

is a result of cancellations. Indeed, Figure 5 shows that the function $D_N(x)$ takes on positive and negative values and oscillates very rapidly as N gets large.

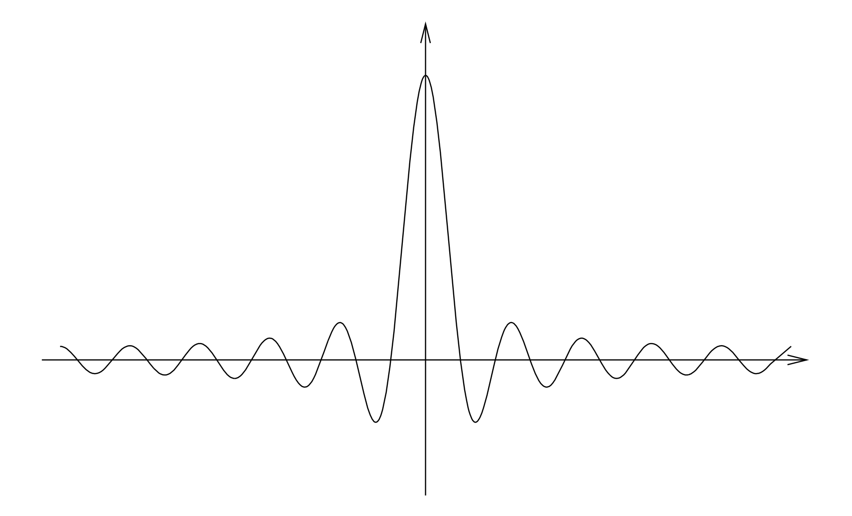


Figure 5. The Dirichlet kernel for large N

This observation suggests that the pointwise convergence of Fourier series is intricate, and may even fail at points of continuity. This is indeed the case, as we will see in the next chapter.

5 Cesàro and Abel summability: applications to Fourier series

Since a Fourier series may fail to converge at individual points, we are led to try to overcome this failure by interpreting the limit

$$\lim_{N\to\infty} S_N(f) = f$$

in a different sense.

5.1 Cesàro means and summation

We begin by taking ordinary averages of the partial sums, a technique which we now describe in more detail.

Suppose we are given a series of complex numbers

$$c_0 + c_1 + c_2 + \cdots = \sum_{k=0}^{\infty} c_k.$$

We define the nth partial sum s_n by

$$s_n = \sum_{k=0}^{n} c_k,$$

and say that the series converges to s if $\lim_{n\to\infty} s_n = s$. This is the most natural and most commonly used type of "summability." Consider, however, the example of the series

$$1 - 1 + 1 - 1 + \cdots = \sum_{k=0}^{\infty} (-1)^k. \tag{3}$$

Its partial sums form the sequence $\{1, 0, 1, 0, \ldots\}$ which has no limit. Because these partial sums alternate evenly between 1 and 0, one might therefore suggest that 1/2 is the "limit" of the sequence, and hence 1/2 equals the "sum" of that particular series. We give a precise meaning to this by defining the average of the first N partial sums by

$$\sigma_N = \frac{s_0 + s_1 + \cdots + s_{N-1}}{N}.$$

The quantity σ_N is called the Nth **Cesàro mean**[7] of the sequence $\{s_k\}$ or the Nth **Cesàro sum** of the series $\sum_{k=0}^{\infty} c_k$.

If σ_N converges to a limit σ as N tends to infinity, we say that the series $\sum c_n$ is **Cesàro summable** to σ. In the case of series of functions, we shall understand the limit in the sense of either pointwise or uniform convergence, depending on the situation.

The reader will have no difficulty checking that in the above example (3), the series is Cesàro summable to 1/2. Moreover, one can show that Cesàro summation is a more inclusive process than convergence. In fact, if a series is convergent to s, then it is also Cesàro summable to the same limit s (Exercise 12).

5.2 Fejér's theorem

An interesting application of Cesàro summability appears in the context of Fourier series.

[7]Note that if the series $\sum_{k=1}^{\infty} c_k$ begins with the term $k = 1$, then it is common practice to define $\sigma_N = (s_1 + \cdots + s_N)/N$. This change of notation has little effect on what follows.

We mentioned earlier that the Dirichlet kernels fail to belong to the family of good kernels. Quite surprisingly, their averages are very well behaved functions, in the sense that they do form a family of good kernels.

To see this, we form the Nth Cesàro mean of the Fourier series, which by definition is

$$\sigma_N(f)(x) = \frac{S_0(f)(x) + \cdots + S_{N-1}(f)(x)}{N}.$$

Since $S_n(f) = f * D_n$, we find that

$$\sigma_N(f)(x) = (f * F_N)(x),$$

where $F_N(x)$ is the N-th **Fejér kernel** given by

$$F_N(x) = \frac{D_0(x) + \cdots + D_{N-1}(x)}{N}.$$

Lemma 5.1 *We have*

$$F_N(x) = \frac{1}{N}\frac{\sin^2(Nx/2)}{\sin^2(x/2)},$$

and the Fejér kernel is a good kernel.

The proof of the formula for F_N (a simple application of trigonometric identities) is outlined in Exercise 15. To prove the rest of the lemma, note that F_N is positive and $\frac{1}{2\pi}\int_{-\pi}^{\pi} F_N(x)\,dx = 1$, in view of the fact that a similar identity holds for the Dirichlet kernels D_n. However, $\sin^2(x/2) \geq c_\delta > 0$, if $\delta \leq |x| \leq \pi$, hence $F_N(x) \leq 1/(Nc_\delta)$, from which it follows that

$$\int_{\delta \leq |x| \leq \pi} |F_N(x)|\,dx \to 0 \quad \text{as } N \to \infty.$$

Applying Theorem 4.1 to this new family of good kernels yields the following important result.

Theorem 5.2 *If f is integrable on the circle, then the Fourier series of f is Cesàro summable to f at every point of continuity of f.*

Moreover, if f is continuous on the circle, then the Fourier series of f is uniformly Cesàro summable to f.

We may now state two corollaries. The first is a result that we have already established. The second is new, and of fundamental importance.

Corollary 5.3 *If f is integrable on the circle and $\hat{f}(n) = 0$ for all n, then $f = 0$ at all points of continuity of f.*

The proof is immediate since all the partial sums are 0, hence all the Cesàro means are 0.

Corollary 5.4 *Continuous functions on the circle can be uniformly approximated by trigonometric polynomials.*

This means that if f is continuous on $[-\pi, \pi]$ with $f(-\pi) = f(\pi)$ and $\epsilon > 0$, then there exists a trigonometric polynomial P such that

$$|f(x) - P(x)| < \epsilon \quad \text{for all } -\pi \leq x \leq \pi.$$

This follows immediately from the theorem since the partial sums, hence the Cesàro means, are trigonometric polynomials. Corollary 5.4 is the periodic analogue of the Weierstrass approximation theorem for polynomials which can be found in Exercise 16.

5.3 Abel means and summation

Another method of summation was first considered by Abel and actually predates the Cesàro method.

A series of complex numbers $\sum_{k=0}^{\infty} c_k$ is said to be **Abel summable** to s if for every $0 \leq r < 1$, the series

$$A(r) = \sum_{k=0}^{\infty} c_k r^k$$

converges, and

$$\lim_{r \to 1} A(r) = s.$$

The quantities $A(r)$ are called the **Abel means** of the series. One can prove that if the series converges to s, then it is Abel summable to s. Moreover, the method of Abel summability is even more powerful than the Cesàro method: when the series is Cesàro summable, it is always Abel summable to the same sum. However, if we consider the series

$$1 - 2 + 3 - 4 + 5 - \cdots = \sum_{k=0}^{\infty} (-1)^k (k+1),$$

then one can show that it is Abel summable to $1/4$ since

$$A(r) = \sum_{k=0}^{\infty} (-1)^k (k+1) r^k = \frac{1}{(1+r)^2},$$

but this series is not Cesàro summable; see Exercise 13.

5.4 The Poisson kernel and Dirichlet's problem in the unit disc

To adapt Abel summability to the context of Fourier series, we define the Abel means of the function $f(\theta) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ by

$$A_r(f)(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} a_n e^{in\theta}.$$

Since the index n takes positive and negative values, it is natural to write $c_0 = a_0$, and $c_n = a_n e^{in\theta} + a_{-n} e^{-in\theta}$ for $n > 0$, so that the Abel means of the Fourier series correspond to the definition given in the previous section for numerical series.

We note that since f is integrable, $|a_n|$ is uniformly bounded in n, so that $A_r(f)$ converges absolutely and uniformly for each $0 \le r < 1$. Just as in the case of Cesàro means, the key fact is that these Abel means can be written as convolutions

$$A_r(f)(\theta) = (f * P_r)(\theta),$$

where $P_r(\theta)$ is the **Poisson kernel** given by

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}. \tag{4}$$

In fact,

$$\begin{aligned} A_r(f)(\theta) &= \sum_{n=-\infty}^{\infty} r^{|n|} a_n e^{in\theta} \\ &= \sum_{n=-\infty}^{\infty} r^{|n|} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\varphi) e^{-in\varphi}\, d\varphi \right) e^{in\theta} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\varphi) \left(\sum_{n=-\infty}^{\infty} r^{|n|} e^{-in(\varphi-\theta)} \right) d\varphi, \end{aligned}$$

where the interchange of the integral and infinite sum is justified by the uniform convergence of the series.

Lemma 5.5 *If $0 \le r < 1$, then*

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}.$$

The Poisson kernel is a good kernel,[8] *as r tends to 1 from below.*

Proof. The identity $P_r(\theta) = \frac{1-r^2}{1-2r\cos\theta+r^2}$ has already been derived in Section 1.1. Note that

$$1 - 2r\cos\theta + r^2 = (1-r)^2 + 2r(1-\cos\theta).$$

Hence if $1/2 \leq r \leq 1$ and $\delta \leq |\theta| \leq \pi$, then

$$1 - 2r\cos\theta + r^2 \geq c_\delta > 0.$$

Thus $P_r(\theta) \leq (1-r^2)/c_\delta$ when $\delta \leq |\theta| \leq \pi$, and the third property of good kernels is verified. Clearly $P_r(\theta) \geq 0$, and integrating the expression (4) term by term (which is justified by the absolute convergence of the series) yields

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta)\, d\theta = 1,$$

thereby concluding the proof that P_r is a good kernel.

Combining this lemma with Theorem 4.1, we obtain our next result.

Theorem 5.6 *The Fourier series of an integrable function on the circle is Abel summable to f at every point of continuity. Moreover, if f is continuous on the circle, then the Fourier series of f is uniformly Abel summable to f.*

We now return to a problem discussed in Chapter 1, where we sketched the solution of the steady-state heat equation $\triangle u = 0$ in the unit disc with boundary condition $u = f$ on the circle. We expressed the Laplacian in terms of polar coordinates, separated variables, and expected that a solution was given by

$$u(r,\theta) = \sum_{m=-\infty}^{\infty} a_m r^{|m|} e^{im\theta}, \tag{5}$$

where a_m was the mth Fourier coefficient of f. In other words, we were led to take

$$u(r,\theta) = A_r(f)(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\varphi)P_r(\theta-\varphi)\, d\varphi.$$

We are now in a position to show that this is indeed the case.

[8]In this case, the family of kernels is indexed by a continuous parameter $0 \leq r < 1$, rather than the discrete n considered previously. In the definition of good kernels, we simply replace n by r and take the limit in property (c) appropriately, for example $r \to 1$ in this case.

Theorem 5.7 *Let f be an integrable function defined on the unit circle. Then the function u defined in the unit disc by the Poisson integral*

$$u(r,\theta) = (f * P_r)(\theta) \tag{6}$$

has the following properties:

(i) *u has two continuous derivatives in the unit disc and satisfies $\triangle u = 0$.*

(ii) *If θ is any point of continuity of f, then*

$$\lim_{r \to 1} u(r,\theta) = f(\theta).$$

If f is continuous everywhere, then this limit is uniform.

(iii) *If f is continuous, then $u(r,\theta)$ is the unique solution to the steady-state heat equation in the disc which satisfies conditions* (i) *and* (ii).

Proof. For (i), we recall that the function u is given by the series (5). Fix $\rho < 1$; inside each disc of radius $r < \rho < 1$ centered at the origin, the series for u can be differentiated term by term, and the differentiated series is uniformly and absolutely convergent. Thus u can be differentiated twice (in fact infinitely many times), and since this holds for all $\rho < 1$, we conclude that u is twice differentiable inside the unit disc. Moreover, in polar coordinates,

$$\triangle u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2},$$

so term by term differentiation shows that $\triangle u = 0$.

The proof of (ii) is a simple application of the previous theorem. To prove (iii) we argue as follows. Suppose v solves the steady-state heat equation in the disc and converges to f uniformly as r tends to 1 from below. For each fixed r with $0 < r < 1$, the function $v(r,\theta)$ has a Fourier series

$$\sum_{n=-\infty}^{\infty} a_n(r) e^{in\theta} \quad \text{where} \quad a_n(r) = \frac{1}{2\pi}\int_{-\pi}^{\pi} v(r,\theta) e^{-in\theta}\, d\theta.$$

Taking into account that $v(r,\theta)$ solves the equation

$$\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial^2 v}{\partial \theta^2} = 0, \tag{7}$$

we find that

$$a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{r^2}a_n(r) = 0. \tag{8}$$

Indeed, we may first multiply (7) by $e^{-in\theta}$ and integrate in θ. Then, since v is periodic, two integrations by parts give

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\partial^2 v}{\partial \theta^2}(r,\theta)e^{-in\theta}\, d\theta = -n^2 a_n(r).$$

Finally, we may interchange the order of differentiation and integration, which is permissible since v has two continuous derivatives; this yields (8).

Therefore, we must have $a_n(r) = A_n r^n + B_n r^{-n}$ for some constants A_n and B_n, when $n \neq 0$ (see Exercise 11 in Chapter 1). To evaluate the constants, we first observe that each term $a_n(r)$ is bounded because v is bounded, therefore $B_n = 0$. To find A_n we let $r \to 1$. Since v converges uniformly to f as $r \to 1$ we find that

$$A_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}\, d\theta.$$

By a similar argument, this formula also holds when $n = 0$. Our conclusion is that for each $0 < r < 1$, the Fourier series of v is given by the series of $u(r,\theta)$, so by the uniqueness of Fourier series for continuous functions, we must have $u = v$.

Remark. By part (iii) of the theorem, we may conclude that if u solves $\triangle u = 0$ in the disc, and converges to 0 uniformly as $r \to 1$, then u must be identically 0. However, if uniform convergence is replaced by pointwise convergence, this conclusion may fail; see Exercise 18.

6 Exercises

1. Suppose f is 2π-periodic and integrable on any finite interval. Prove that if $a, b \in \mathbb{R}$, then

$$\int_a^b f(x)\, dx = \int_{a+2\pi}^{b+2\pi} f(x)\, dx = \int_{a-2\pi}^{b-2\pi} f(x)\, dx.$$

Also prove that

$$\int_{-\pi}^{\pi} f(x+a)\, dx = \int_{-\pi}^{\pi} f(x)\, dx = \int_{-\pi+a}^{\pi+a} f(x)\, dx.$$

2. In this exercise we show how the symmetries of a function imply certain properties of its Fourier coefficients. Let f be a 2π-periodic Riemann integrable function defined on $\mathbb{R}$.

(a) Show that the Fourier series of the function f can be written as

$$f(\theta) \sim \hat{f}(0) + \sum_{n\geq 1}[\hat{f}(n) + \hat{f}(-n)]\cos n\theta + i[\hat{f}(n) - \hat{f}(-n)]\sin n\theta.$$

(b) Prove that if f is even, then $\hat{f}(n) = \hat{f}(-n)$, and we get a cosine series.

(c) Prove that if f is odd, then $\hat{f}(n) = -\hat{f}(-n)$, and we get a sine series.

(d) Suppose that $f(\theta + \pi) = f(\theta)$ for all $\theta \in \mathbb{R}$. Show that $\hat{f}(n) = 0$ for all odd n.

(e) Show that f is real-valued if and only if $\overline{\hat{f}(n)} = \hat{f}(-n)$ for all n.

3. We return to the problem of the plucked string discussed in Chapter 1. Show that the initial condition f is *equal* to its Fourier sine series

$$f(x) = \sum_{m=1}^{\infty} A_m \sin mx \qquad \text{with} \qquad A_m = \frac{2h}{m^2}\frac{\sin mp}{p(\pi - p)}.$$

[Hint: Note that $|A_m| \leq C/m^2$.]

4. Consider the 2π-periodic odd function defined on $[0, \pi]$ by $f(\theta) = \theta(\pi - \theta)$.

(a) Draw the graph of f.

(b) Compute the Fourier coefficients of f, and show that

$$f(\theta) = \frac{8}{\pi} \sum_{k \text{ odd } \geq 1} \frac{\sin k\theta}{k^3}.$$

5. On the interval $[-\pi, \pi]$ consider the function

$$f(\theta) = \begin{cases} 0 & \text{if } |\theta| > \delta, \\ 1 - |\theta|/\delta & \text{if } |\theta| \leq \delta. \end{cases}$$

Thus the graph of f has the shape of a triangular tent. Show that

$$f(\theta) = \frac{\delta}{2\pi} + 2\sum_{n=1}^{\infty} \frac{1 - \cos n\delta}{n^2\pi\delta}\cos n\theta.$$

6. Let f be the function defined on $[-\pi, \pi]$ by $f(\theta) = |\theta|$.

(a) Draw the graph of f.

(b) Calculate the Fourier coefficients of f, and show that

$$\hat{f}(n) = \begin{cases} \dfrac{\pi}{2} & \text{if } n = 0, \\ \dfrac{-1 + (-1)^n}{\pi n^2} & \text{if } n \neq 0. \end{cases}$$

(c) What is the Fourier series of f in terms of sines and cosines?

(d) Taking $\theta = 0$, prove that

$$\sum_{n \text{ odd } \geq 1} \frac{1}{n^2} = \frac{\pi^2}{8} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

See also Example 2 in Section 1.1.

7. Suppose $\{a_n\}_{n=1}^N$ and $\{b_n\}_{n=1}^N$ are two finite sequences of complex numbers. Let $B_k = \sum_{n=1}^k b_n$ denote the partial sums of the series $\sum b_n$ with the convention $B_0 = 0$.

(a) Prove the **summation by parts** formula

$$\sum_{n=M}^{N} a_n b_n = a_N B_N - a_M B_{M-1} - \sum_{n=M}^{N-1} (a_{n+1} - a_n) B_n.$$

(b) Deduce from this formula Dirichlet's test for convergence of a series: if the partial sums of the series $\sum b_n$ are bounded, and $\{a_n\}$ is a sequence of real numbers that decreases monotonically to 0, then $\sum a_n b_n$ converges.

8. Verify that $\frac{1}{2i} \sum_{n \neq 0} \frac{e^{inx}}{n}$ is the Fourier series of the 2π-periodic **sawtooth** function illustrated in Figure 6, defined by $f(0) = 0$, and

$$f(x) = \begin{cases} -\dfrac{\pi}{2} - \dfrac{x}{2} & \text{if } -\pi < x < 0, \\ \dfrac{\pi}{2} - \dfrac{x}{2} & \text{if } 0 < x < \pi. \end{cases}$$

Note that this function is not continuous. Show that nevertheless, the series converges for every x (by which we mean, as usual, that the symmetric partial sums of the series converge). In particular, the value of the series at the origin, namely 0, is the average of the values of $f(x)$ as x approaches the origin from the left and the right.

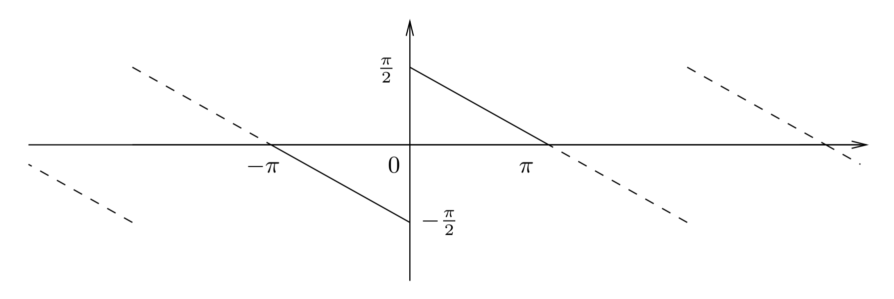


Figure 6. The sawtooth function

[Hint: Use Dirichlet's test for convergence of a series $\sum a_n b_n$.]

9. Let $f(x) = \chi_{[a,b]}(x)$ be the characteristic function of the interval $[a, b] \subset [-\pi, \pi]$, that is,

$$\chi_{[a,b]}(x) = \begin{cases} 1 & \text{if } x \in [a, b], \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that the Fourier series of f is given by

$$f(x) \sim \frac{b-a}{2\pi} + \sum_{n \neq 0} \frac{e^{-ina} - e^{-inb}}{2\pi in} e^{inx}.$$

The sum extends over all positive and negative integers excluding 0.

(b) Show that if $a \neq -\pi$ or $b \neq \pi$ and $a \neq b$, then the Fourier series does not converge absolutely for any x. [Hint: It suffices to prove that for many values of n one has $|\sin n\theta_0| \geq c > 0$ where $\theta_0 = (b-a)/2$.]

(c) However, prove that the Fourier series converges at every point x. What happens if $a = -\pi$ and $b = \pi$?

10. Suppose f is a periodic function of period 2π which belongs to the class C^k. Show that

$$\hat{f}(n) = O(1/|n|^k) \quad \text{as } |n| \to \infty.$$

This notation means that there exists a constant C such $|\hat{f}(n)| \leq C/|n|^k$. We could also write this as $|n|^k \hat{f}(n) = O(1)$, where $O(1)$ means bounded.

[Hint: Integrate by parts.]

11. Suppose that $\{f_k\}_{k=1}^{\infty}$ is a sequence of Riemann integrable functions on the interval $[0, 1]$ such that

$$\int_0^1 |f_k(x) - f(x)|\, dx \to 0 \quad \text{as } k \to \infty.$$

Show that $\hat{f}_k(n) \to \hat{f}(n)$ uniformly in n as $k \to \infty$.

12. Prove that if a series of complex numbers $\sum c_n$ converges to s, then $\sum c_n$ is Cesàro summable to s.
[Hint: Assume $s_n \to 0$ as $n \to \infty$.]

13. The purpose of this exercise is to prove that Abel summability is stronger than the standard or Cesàro methods of summation.

(a) Show that if the series $\sum_{n=1}^{\infty} c_n$ of complex numbers converges to a finite limit s, then the series is Abel summable to s. [Hint: Why is it enough to prove the theorem when $s = 0$? Assuming $s = 0$, show that if $s_N = c_1 + \cdots + c_N$, then $\sum_{n=1}^{N} c_n r^n = (1-r)\sum_{n=1}^{N} s_n r^n + s_N r^{N+1}$. Let $N \to \infty$ to show that

$$\sum c_n r^n = (1-r)\sum s_n r^n.$$

Finally, prove that the right-hand side converges to 0 as $r \to 1$.]

(b) However, show that there exist series which are Abel summable, but that do not converge. [Hint: Try $c_n = (-1)^n$. What is the Abel limit of $\sum c_n$?]

(c) Argue similarly to prove that if a series $\sum_{n=1}^{\infty} c_n$ is Cesàro summable to σ, then it is Abel summable to σ. [Hint: Note that

$$\sum_{n=1}^{\infty} c_n r^n = (1-r)^2 \sum_{n=1}^{\infty} n\sigma_n r^n,$$

and assume $\sigma = 0$.]

(d) Give an example of a series that is Abel summable but not Cesàro summable. [Hint: Try $c_n = (-1)^{n-1} n$. Note that if $\sum c_n$ is Cesàro summable, then c_n/n tends to 0.]

The results above can be summarized by the following implications about series:

$$\text{convergent} \Longrightarrow \text{Cesàro summable} \Longrightarrow \text{Abel summable},$$

and the fact that none of the arrows can be reversed.

14. This exercise deals with a theorem of Tauber which says that under an additional condition on the coefficients c_n, the above arrows can be reversed.

(a) If $\sum c_n$ is Cesàro summable to σ and $c_n = o(1/n)$ (that is, $nc_n \to 0$), then $\sum c_n$ converges to σ. [Hint: $s_n - \sigma_n = [(n-1)c_n + \cdots + c_2]/n$.]

(b) The above statement holds if we replace Cesàro summable by Abel summable. [Hint: Estimate the difference between $\sum_{n=1}^{N} c_n$ and $\sum_{n=1}^{N} c_n r^n$ where $r = 1 - 1/N$.]

15. Prove that the Fejér kernel is given by

$$F_N(x) = \frac{1}{N} \frac{\sin^2(Nx/2)}{\sin^2(x/2)}.$$

[Hint: Remember that $NF_N(x) = D_0(x) + \cdots + D_{N-1}(x)$ where $D_n(x)$ is the Dirichlet kernel. Therefore, if $\omega = e^{ix}$ we have

$$NF_N(x) = \sum_{n=0}^{N-1} \frac{\omega^{-n} - \omega^{n+1}}{1-\omega}.]$$

16. The Weierstrass approximation theorem states: Let f be a continuous function on the closed and bounded interval $[a,b] \subset \mathbb{R}$. Then, for any $\epsilon > 0$, there exists a polynomial P such that

$$\sup_{x\in[a,b]} |f(x) - P(x)| < \epsilon.$$

Prove this by applying Corollary 5.4 of Fejér's theorem and using the fact that the exponential function e^{ix} can be approximated by polynomials uniformly on any interval.

17. In Section 5.4 we proved that the Abel means of f converge to f at all points of continuity, that is,

$$\lim_{r\to 1} A_r(f)(\theta) = \lim_{r\to 1} (P_r * f)(\theta) = f(\theta), \qquad \text{with } 0 < r < 1,$$

whenever f is continuous at θ. In this exercise, we will study the behavior of $A_r(f)(\theta)$ at certain points of discontinuity.

An integrable function is said to have a **jump discontinuity** at θ if the two limits

$$\lim_{\substack{h\to 0\\ h>0}} f(\theta+h) = f(\theta^+) \qquad \text{and} \qquad \lim_{\substack{h\to 0\\ h>0}} f(\theta-h) = f(\theta^-)$$

exist.

(a) Prove that if f has a jump discontinuity at θ, then

$$\lim_{r\to 1} A_r(f)(\theta) = \frac{f(\theta^+) + f(\theta^-)}{2}, \qquad \text{with } 0 \le r < 1.$$

[Hint: Explain why $\frac{1}{2\pi}\int_{-\pi}^{0} P_r(\theta)\, d\theta = \frac{1}{2\pi}\int_0^{\pi} P_r(\theta)\, d\theta = \frac{1}{2}$, then modify the proof given in the text.]

(b) Using a similar argument, show that if f has a jump discontinuity at θ, the Fourier series of f at θ is Cesàro summable to $\frac{f(\theta^+)+f(\theta^-)}{2}$.

18. If $P_r(\theta)$ denotes the Poisson kernel, show that the function

$$u(r,\theta) = \frac{\partial P_r}{\partial \theta},$$

defined for $0 \le r < 1$ and $\theta \in \mathbb{R}$, satisfies:

(i) $\triangle u = 0$ in the disc.

(ii) $\lim_{r\to 1} u(r,\theta) = 0$ for each θ.

However, u is not identically zero.

19. Solve Laplace's equation $\triangle u = 0$ in the semi infinite strip

$$S = \{(x,y) : 0 < x < 1,\ 0 < y\},$$

subject to the following boundary conditions

$$\begin{cases} u(0,y) = 0 & \text{when } 0 \le y, \\ u(1,y) = 0 & \text{when } 0 \le y, \\ u(x,0) = f(x) & \text{when } 0 \le x \le 1 \end{cases}$$

where f is a given function, with of course $f(0) = f(1) = 0$. Write

$$f(x) = \sum_{n=1}^{\infty} a_n \sin(n\pi x)$$

and expand the general solution in terms of the special solutions given by

$$u_n(x,y) = e^{-n\pi y}\sin(n\pi x).$$

Express u as an integral involving f, analogous to the Poisson integral formula (6).

20. Consider the Dirichlet problem in the annulus defined by $\{(r,\theta) : \rho < r < 1\}$, where $0 < \rho < 1$ is the inner radius. The problem is to solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0$$

subject to the boundary conditions

$$\begin{cases} u(1,\theta) = f(\theta), \\ u(\rho,\theta) = g(\theta), \end{cases}$$

where f and g are given continuous functions.

Arguing as we have previously for the Dirichlet problem in the disc, we can hope to write

$$u(r,\theta) = \sum c_n(r)e^{in\theta}$$

with $c_n(r) = A_n r^n + B_n r^{-n}$, $n \neq 0$. Set

$$f(\theta) \sim \sum a_n e^{in\theta} \quad \text{and} \quad g(\theta) \sim \sum b_n e^{in\theta}.$$

We want $c_n(1) = a_n$ and $c_n(\rho) = b_n$. This leads to the solution

$$u(r,\theta) = \sum_{n\neq 0} \left(\frac{1}{\rho^n - \rho^{-n}}\right) \left[((\rho/r)^n - (r/\rho)^n)\, a_n + (r^n - r^{-n})b_n\right] e^{in\theta}$$
$$+a_0 + (b_0 - a_0)\frac{\log r}{\log \rho}.$$

Show that as a result we have

$$u(r,\theta) - (P_r * f)(\theta) \to 0 \quad \text{as } r \to 1 \text{ uniformly in } \theta,$$

and

$$u(r,\theta) - (P_{\rho/r} * g)(\theta) \to 0 \quad \text{as } r \to \rho \text{ uniformly in } \theta.$$

7 Problems

1. One can construct Riemann integrable functions on $[0,1]$ that have a dense set of discontinuities as follows.

(a) Let $f(x) = 0$ when $x < 0$, and $f(x) = 1$ if $x \geq 0$. Choose a countable dense sequence $\{r_n\}$ in $[0,1]$. Then, show that the function

$$F(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} f(x - r_n)$$

is integrable and has discontinuities at all points of the sequence $\{r_n\}$. [Hint: F is monotonic and bounded.]

(b) Consider next

$$F(x) = \sum_{n=1}^{\infty} 3^{-n} g(x - r_n),$$

where $g(x) = \sin 1/x$ when $x \neq 0$, and $g(0) = 0$. Then F is integrable, discontinuous at each $x = r_n$, and fails to be monotonic in any subinterval of $[0,1]$. [Hint: Use the fact that $3^{-k} > \sum_{n>k} 3^{-n}$.]

(c) The original example of Riemann is the function

$$F(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2},$$

where $(x) = x$ for $x \in (-1/2, 1/2]$ and (x) is continued to $\mathbb{R}$ by periodicity, that is, $(x+1) = (x)$. It can be shown that F is discontinuous whenever $x = m/2n$, where $m, n \in \mathbb{Z}$ with m odd and $n \neq 0$.

2. Let D_N denote the Dirichlet kernel

$$D_N(\theta) = \sum_{k=-N}^{N} e^{ik\theta} = \frac{\sin((N+1/2)\theta)}{\sin(\theta/2)},$$

and define

$$L_N = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)|\, d\theta.$$

(a) Prove that

$$L_N \geq c \log N$$

for some constant $c > 0$. [Hint: Show that $|D_N(\theta)| \geq c \frac{\sin((N+1/2)\theta)}{|\theta|}$, change variables, and prove that

$$L_N \geq c \int_{\pi}^{N\pi} \frac{|\sin\theta|}{|\theta|}\, d\theta + O(1).$$

Write the integral as a sum $\sum_{k=1}^{N-1} \int_{k\pi}^{(k+1)\pi}$. To conclude, use the fact that $\sum_{k=1}^{n} 1/k \geq c \log n$.] A more careful estimate gives

$$L_N = \frac{4}{\pi^2} \log N + O(1).$$

(b) Prove the following as a consequence: for each $n \geq 1$, there exists a continuous function f_n such that $|f_n| \leq 1$ and $|S_n(f_n)(0)| \geq c' \log n$. [Hint: The function g_n which is equal to 1 when D_n is positive and -1 when D_n is negative has the desired property but is not continuous. Approximate g_n in the integral norm (in the sense of Lemma 3.2) by continuous functions h_k satisfying $|h_k| \leq 1$.]

3.* Littlewood provided a refinement of Tauber's theorem:

(a) If $\sum c_n$ is Abel summable to s and $c_n = O(1/n)$, then $\sum c_n$ converges to s.

(b) As a consequence, if $\sum c_n$ is Cesàro summable to s and $c_n = O(1/n)$, then $\sum c_n$ converges to s.

These results may be applied to Fourier series. By Exercise 17, they imply that if f is an integrable function that satisfies $\hat{f}(\nu) = O(1/|\nu|)$, then:

(i) If f is continuous at θ, then

$$S_N(f)(\theta) \to f(\theta) \qquad \text{as } N \to \infty.$$

(ii) If f has a jump discontinuity at θ, then

$$S_N(f)(\theta) \to \frac{f(\theta^+) + f(\theta^-)}{2} \qquad \text{as } N \to \infty.$$

(iii) If f is continuous on $[-\pi, \pi]$, then $S_N(f) \to f$ uniformly.

For the simpler assertion (b), hence a proof of (i), (ii), and (iii), see Problem 5 in Chapter 4.

3 Convergence of Fourier Series

> The sine and cosine series, by which one can represent an arbitrary function in a given interval, enjoy among other remarkable properties that of being convergent. This property did not escape the great geometer (Fourier) who began, through the introduction of the representation of functions just mentioned, a new career for the applications of analysis; it was stated in the Memoir which contains his first research on heat. But no one so far, to my knowledge, gave a general proof of it . . .
>
> *G. Dirichlet,* 1829

In this chapter, we continue our study of the problem of convergence of Fourier series. We approach the problem from two different points of view.

The first is "global" and concerns the overall behavior of a function f over the entire interval $[0, 2\pi]$. The result we have in mind is "mean-square convergence": if f is integrable on the circle, then

$$\frac{1}{2\pi}\int_0^{2\pi} |f(\theta) - S_N(f)(\theta)|^2\, d\theta \to 0 \quad \text{as } N \to \infty.$$

At the heart of this result is the fundamental notion of "orthogonality"; this idea is expressed in terms of vector spaces with inner products, and their related infinite dimensional variants, the Hilbert spaces. A connected result is the Parseval identity which equates the mean-square "norm" of the function with a corresponding norm of its Fourier coefficients. Orthogonality is a fundamental mathematical notion which has many applications in analysis.

The second viewpoint is "local" and concerns the behavior of f near a given point. The main question we consider is the problem of pointwise convergence: does the Fourier series of f converge to the value $f(\theta)$ for a given θ? We first show that this convergence does indeed hold whenever f is differentiable at θ. As a corollary, we obtain the Riemann localization principle, which states that the question of whether or not $S_N(f)(\theta) \to f(\theta)$ is completely determined by the behavior of f in an

arbitrarily small interval about θ. This is a remarkable result since the Fourier coefficients, hence the Fourier series, of f depend on the values of f on the whole interval $[0, 2\pi]$.

Even though convergence of the Fourier series holds at points where f is differentiable, it may fail if f is merely continuous. The chapter concludes with the presentation of a continuous function whose Fourier series does not converge at a given point, as promised earlier.

1 Mean-square convergence of Fourier series

The aim of this section is the proof of the following theorem.

Theorem 1.1 *Suppose f is integrable on the circle. Then*

$$\frac{1}{2\pi}\int_0^{2\pi} |f(\theta) - S_N(f)(\theta)|^2 \, d\theta \to 0 \qquad \text{as } N \to \infty.$$

As we remarked earlier, the key concept involved is that of orthogonality. The correct setting for orthogonality is in a vector space equipped with an inner product.

1.1 Vector spaces and inner products

We now review the definitions of a vector space over $\mathbb{R}$ or $\mathbb{C}$, an inner product, and its associated norm. In addition to the familiar finite-dimensional vector spaces $\mathbb{R}^d$ and $\mathbb{C}^d$, we also examine two infinite-dimensional examples which play a central role in the proof of Theorem 1.1.

Preliminaries on vector spaces

A vector space V over the real numbers $\mathbb{R}$ is a set whose elements may be "added" together, and "multiplied" by scalars. More precisely, we may associate to any pair $X, Y \in V$ an element in V called their sum and denoted by $X + Y$. We require that this addition respects the usual laws of arithmetic, such as commutativity $X + Y = Y + X$, and associativity $X + (Y + Z) = (X + Y) + Z$, etc. Also, given any $X \in V$ and real number λ, we assign an element $\lambda X \in V$ called the product of X by λ. This scalar multiplication must satisfy the standard properties, for instance $\lambda_1(\lambda_2 X) = (\lambda_1\lambda_2)X$ and $\lambda(X + Y) = \lambda X + \lambda Y$. We may instead allow scalar multiplication by numbers in $\mathbb{C}$; we then say that V is a vector space over the complex numbers.

For example, the set $\mathbb{R}^d$ of d-tuples of real numbers $(x_1, x_2, \dots, x_d)$ is a vector space over the reals. Addition is defined componentwise by

$$(x_1, \dots, x_d) + (y_1, \dots, y_d) = (x_1 + y_1, \dots, x_d + y_d),$$

and so is multiplication by a scalar $\lambda \in \mathbb{R}$:

$$\lambda(x_1, \dots, x_d) = (\lambda x_1, \dots, \lambda x_d).$$

Similarly, the space $\mathbb{C}^d$ (the complex version of the previous example) is the set of d-tuples of complex numbers $(z_1, z_2, \dots, z_d)$. It is a vector space over $\mathbb{C}$ with addition defined componentwise by

$$(z_1, \dots, z_d) + (w_1, \dots, w_d) = (z_1 + w_1, \dots, z_d + w_d).$$

Multiplication by scalars $\lambda \in \mathbb{C}$ is given by

$$\lambda(z_1, \dots, z_d) = (\lambda z_1, \dots, \lambda z_d).$$

An **inner product** on a vector space V over $\mathbb{R}$ associates to any pair X, Y of elements in V a real number which we denote by (X, Y). In particular, the inner product must be symmetric $(X, Y) = (Y, X)$ and linear in both variables; that is,

$$(\alpha X + \beta Y, Z) = \alpha(X, Z) + \beta(Y, Z)$$

whenever $\alpha, \beta \in \mathbb{R}$ and $X, Y, Z \in V$. Also, we require that the inner product be positive-definite, that is, $(X, X) \geq 0$ for all X in V. In particular, given an inner product $(\cdot, \cdot)$ we may define the norm of X by

$$\|X\| = (X, X)^{1/2}.$$

If in addition $\|X\| = 0$ implies $X = 0$, we say that the inner product is strictly positive-definite.

For example, the space $\mathbb{R}^d$ is equipped with a (strictly positive-definite) inner product defined by

$$(X, Y) = x_1 y_1 + \cdots + x_d y_d$$

when $X = (x_1, \dots, x_d)$ and $Y = (y_1, \dots, y_d)$. Then

$$\|X\| = (X, X)^{1/2} = \sqrt{x_1^2 + \cdots + x_d^2},$$

which is the usual Euclidean distance. One also uses the notation $|X|$ instead of $\|X\|$.

For vector spaces over the complex numbers, the inner product of two elements is a complex number. Moreover, these inner products are called Hermitian (instead of symmetric) since they must satisfy $(X, Y) = \overline{(Y, X)}$. Hence the inner product is linear in the first variable, but conjugate-linear in the second:

$$(\alpha X + \beta Y, Z) = \alpha(X, Z) + \beta(Y, Z) \quad \text{and}$$

$$(X, \alpha Y + \beta Z) = \overline{\alpha}(X, Y) + \overline{\beta}(X, Z).$$

Also, we must have $(X, X) \geq 0$, and the norm of X is defined by $\|X\| = (X, X)^{1/2}$ as before. Again, the inner product is strictly positive-definite if $\|X\| = 0$ implies $X = 0$.

For example, the inner product of two vectors $Z = (z_1, \ldots, z_d)$ and $W = (w_1, \ldots, w_d)$ in $\mathbb{C}^d$ is defined by

$$(Z, W) = z_1\overline{w_1} + \cdots + z_d\overline{w_d}.$$

The norm of the vector Z is then given by

$$\|Z\| = (Z, Z)^{1/2} = \sqrt{|z_1|^2 + \cdots + |z_d|^2}.$$

The presence of an inner product on a vector space allows one to define the geometric notion of "orthogonality." Let V be a vector space (over $\mathbb{R}$ or $\mathbb{C}$) with inner product $(\cdot, \cdot)$ and associated norm $\|\cdot\|$. Two elements X and Y are **orthogonal** if $(X, Y) = 0$, and we write $X \perp Y$. Three important results can be derived from this notion of orthogonality:

(i) The Pythagorean theorem: if X and Y are orthogonal, then

$$\|X + Y\|^2 = \|X\|^2 + \|Y\|^2.$$

(ii) The Cauchy-Schwarz inequality: for any $X, Y \in V$ we have

$$|(X, Y)| \leq \|X\| \, \|Y\|.$$

(iii) The triangle inequality: for any $X, Y \in V$ we have

$$\|X + Y\| \leq \|X\| + \|Y\|.$$

The proofs of these facts are simple. For (i) it suffices to expand $(X+Y, X+Y)$ and use the assumption that $(X, Y) = 0$.

For (ii), we first dispose of the case when $\|Y\| = 0$ by showing that this implies $(X, Y) = 0$ for all X. Indeed, for all real t we have

$$0 \leq \|X + tY\|^2 = \|X\|^2 + 2t \operatorname{Re}(X, Y)$$

and $\operatorname{Re}(X, Y) \neq 0$ contradicts the inequality if we take t to be large and positive (or negative). Similarly, by considering $\|X + itY\|^2$, we find that $\operatorname{Im}(X, Y) = 0$.

If $\|Y\| \neq 0$, we may set $c = (X, Y)/(Y, Y)$; then $X - cY$ is orthogonal to Y, and therefore also to cY. If we write $X = X - cY + cY$ and apply the Pythagorean theorem, we get

$$\|X\|^2 = \|X - cY\|^2 + \|cY\|^2 \geq |c|^2 \|Y\|^2.$$

Taking square roots on both sides gives the result. Note that we have equality in the above precisely when $X = cY$.

Finally, for (iii) we first note that

$$\|X + Y\|^2 = (X, X) + (X, Y) + (Y, X) + (Y, Y).$$

But $(X, X) = \|X\|^2$, $(Y, Y) = \|Y\|^2$, and by the Cauchy-Schwarz inequality

$$|(X, Y) + (Y, X)| \leq 2 \|X\| \, \|Y\|,$$

therefore

$$\|X + Y\|^2 \leq \|X\|^2 + 2 \|X\| \, \|Y\| + \|Y\|^2 = (\|X\| + \|Y\|)^2.$$

Two important examples

The vector spaces $\mathbb{R}^d$ and $\mathbb{C}^d$ are finite dimensional. In the context of Fourier series, we need to work with two infinite-dimensional vector spaces, which we now describe.

EXAMPLE 1. The vector space $\ell^2(\mathbb{Z})$ over $\mathbb{C}$ is the set of all (two-sided) infinite sequences of complex numbers

$$(\ldots, a_{-n}, \ldots, a_{-1}, a_0, a_1, \ldots, a_n, \ldots)$$

such that

$$\sum_{n \in \mathbb{Z}} |a_n|^2 < \infty;$$

that is, the series converges. Addition is defined componentwise, and so is scalar multiplication. The inner product between the two vectors $A = (\dots, a_{-1}, a_0, a_1, \dots)$ and $B = (\dots, b_{-1}, b_0, b_1, \dots)$ is defined by the absolutely convergent series

$$(A, B) = \sum_{n\in\mathbb{Z}} a_n \overline{b_n}.$$

The norm of A is then given by

$$\|A\| = (A, A)^{1/2} = \left(\sum_{n\in\mathbb{Z}} |a_n|^2\right)^{1/2}.$$

We must first check that $\ell^2(\mathbb{Z})$ is a vector space. This requires that if A and B are two elements in $\ell^2(\mathbb{Z})$, then so is the vector $A + B$. To see this, for each integer $N > 0$ we let A_N denote the truncated element

$$A_N = (\dots, 0, 0, a_{-N}, \dots, a_{-1}, a_0, a_1, \dots, a_N, 0, 0, \dots),$$

where we have set $a_n = 0$ whenever $|n| > N$. We define the truncated element B_N similarly. Then, by the triangle inequality which holds in a finite dimensional Euclidean space, we have

$$\|A_N + B_N\| \leq \|A_N\| + \|B_N\| \leq \|A\| + \|B\|.$$

Thus

$$\sum_{|n|\leq N} |a_n + b_n|^2 \leq (\|A\| + \|B\|)^2,$$

and letting N tend to infinity gives $\sum_{n\in\mathbb{Z}} |a_n + b_n|^2 < \infty$. It also follows that $\|A + B\| \leq \|A\| + \|B\|$, which is the triangle inequality. The Cauchy-Schwarz inequality, which states that the sum $\sum_{n\in\mathbb{Z}} a_n \overline{b}_n$ converges absolutely and that $|(A, B)| \leq \|A\|\,\|B\|$, can be deduced in the same way from its finite analogue.

In the three examples $\mathbb{R}^d$, $\mathbb{C}^d$, and $\ell^2(\mathbb{Z})$, the vector spaces with their inner products and norms satisfy two important properties:

(i) The inner product is strictly positive-definite, that is, $\|X\| = 0$ implies $X = 0$.

(ii) The vector space is **complete**, which by definition means that every Cauchy sequence in the norm converges to a limit in the vector space.

An inner product space with these two properties is called a **Hilbert space**. We see that $\mathbb{R}^d$ and $\mathbb{C}^d$ are examples of finite-dimensional Hilbert spaces, while $\ell^2(\mathbb{Z})$ is an example of an infinite-dimensional Hilbert space (see Exercises 1 and 2). If either of the conditions above fail, the space is called a **pre-Hilbert space**.

We now give an important example of a pre-Hilbert space where both conditions (i) and (ii) fail.

EXAMPLE 2. Let $\mathcal{R}$ denote the set of complex-valued Riemann integrable functions on $[0, 2\pi]$ (or equivalently, integrable functions on the circle). This is a vector space over $\mathbb{C}$. Addition is defined pointwise by

$$(f+g)(\theta) = f(\theta) + g(\theta).$$

Naturally, multiplication by a scalar $\lambda \in \mathbb{C}$ is given by

$$(\lambda f)(\theta) = \lambda \cdot f(\theta).$$

An inner product is defined on this vector space by

$$(f, g) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)\overline{g(\theta)}\, d\theta. \tag{1}$$

The norm of f is then

$$\|f\| = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|^2\, d\theta\right)^{1/2}.$$

One needs to check that the analogue of the Cauchy-Schwarz and triangle inequalities hold in this example; that is, $|(f, g)| \leq \|f\|\,\|g\|$ and $\|f + g\| \leq \|f\| + \|g\|$. While these facts can be obtained as consequences of the corresponding inequalities in the previous examples, the argument is a little elaborate and we prefer to proceed differently.

We first observe that $2AB \leq (A^2 + B^2)$ for any two real numbers A and B. If we set $A = \lambda^{1/2}|f(\theta)|$ and $B = \lambda^{-1/2}|g(\theta)|$ with $\lambda > 0$, we get

$$|f(\theta)\overline{g(\theta)}| \leq \frac{1}{2}(\lambda|f(\theta)|^2 + \lambda^{-1}|g(\theta)|^2).$$

We then integrate this in θ to obtain

$$|(f, g)| \leq \frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|\,|\overline{g(\theta)}|\, d\theta \leq \frac{1}{2}(\lambda\|f\|^2 + \lambda^{-1}\|g\|^2).$$

Then, put $\lambda = \|g\|/\|f\|$ to get the Cauchy-Schwarz inequality. The triangle inequality is then a simple consequence, as we have seen above.

Of course, in our choice of λ we must assume that $\|f\| \neq 0$ and $\|g\| \neq 0$, which leads us to the following observation.

In $\mathcal{R}$, condition (i) for a Hilbert space fails, since $\|f\| = 0$ implies only that f vanishes at its points of continuity. This is not a very serious problem since in the appendix we show that an integrable function is continuous except for a "negligible" set, so that $\|f\| = 0$ implies that f vanishes except on a set of "measure zero." One can get around the difficulty that f is not identically zero by adopting the convention that such functions are actually the zero function, since for the purpose of integration, f behaves precisely like the zero function.

A more essential difficulty is that the space $\mathcal{R}$ is not complete. One way to see this is to start with the function

$$f(\theta) = \begin{cases} 0 & \text{for } \theta = 0, \\ \log(1/\theta) & \text{for } 0 < \theta \leq 2\pi. \end{cases}$$

Since f is not bounded, it does not belong to the space $\mathcal{R}$. Moreover, the sequence of truncations f_n defined by

$$f_n(\theta) = \begin{cases} 0 & \text{for } 0 \leq \theta \leq 1/n, \\ f(\theta) & \text{for } 1/n < \theta \leq 2\pi \end{cases}$$

can easily be seen to form a Cauchy sequence in $\mathcal{R}$ (see Exercise 5). However, this sequence cannot converge to an element in $\mathcal{R}$, since that limit, if it existed, would have to be f; for another example, see Exercise 7.

This and more complicated examples motivate the search for the completion of $\mathcal{R}$, the class of Riemann integrable functions on $[0, 2\pi]$. The construction and identification of this completion, the Lebesgue class $L^2([0, 2\pi])$, represents an important turning point in the development of analysis (somewhat akin to the much earlier completion of the rationals, that is, the passage from $\mathbb{Q}$ to $\mathbb{R}$). A further discussion of these fundamental ideas will be postponed until Book III, where we take up the Lebesgue theory of integration.

We now turn to the proof of Theorem 1.1.

1.2 Proof of mean-square convergence

Consider the space $\mathcal{R}$ of integrable functions on the circle with inner product

$$(f, g) = \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\overline{g(\theta)}\, d\theta$$

and norm $\|f\|$ defined by

$$\|f\|^2 = (f, f) = \frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|^2\, d\theta.$$

With this notation, we must prove that $\|f - S_N(f)\| \to 0$ as N tends to infinity.

For each integer n, let $e_n(\theta) = e^{in\theta}$, and observe that the family $\{e_n\}_{n\in\mathbb{Z}}$ is **orthonormal**; that is,

$$(e_n, e_m) = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m. \end{cases}$$

Let f be an integrable function on the circle, and let a_n denote its Fourier coefficients. An important observation is that these Fourier coefficients are represented by inner products of f with the elements in the orthonormal set $\{e_n\}_{n\in\mathbb{Z}}$:

$$(f, e_n) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-in\theta}\, d\theta = a_n.$$

In particular, $S_N(f) = \sum_{|n|\leq N} a_n e_n$. Then the orthonormal property of the family $\{e_n\}$ and the fact that $a_n = (f, e_n)$ imply that the difference $f - \sum_{|n|\leq N} a_n e_n$ is orthogonal to e_n for all $|n| \leq N$. Therefore, we must have

$$(f - \sum_{|n|\leq N} a_n e_n) \perp \sum_{|n|\leq N} b_n e_n \tag{2}$$

for any complex numbers b_n. We draw two conclusions from this fact.

First, we can apply the Pythagorean theorem to the decomposition

$$f = f - \sum_{|n|\leq N} a_n e_n + \sum_{|n|\leq N} a_n e_n,$$

where we now choose $b_n = a_n$, to obtain

$$\|f\|^2 = \|f - \sum_{|n|\leq N} a_n e_n\|^2 + \|\sum_{|n|\leq N} a_n e_n\|^2.$$

Since the orthonormal property of the family $\{e_n\}_{n\in\mathbb{Z}}$ implies that

$$\|\sum_{|n|\leq N} a_n e_n\|^2 = \sum_{|n|\leq N} |a_n|^2,$$

we deduce that

$$\|f\|^2 = \|f - S_N(f)\|^2 + \sum_{|n|\leq N} |a_n|^2. \tag{3}$$

The second conclusion we may draw from (2) is the following simple lemma.

Lemma 1.2 (Best approximation) *If f is integrable on the circle with Fourier coefficients a_n, then*

$$\|f - S_N(f)\| \leq \|f - \sum_{|n|\leq N} c_n e_n\|$$

for any complex numbers c_n. Moreover, equality holds precisely when $c_n = a_n$ for all $|n| \leq N$.

Proof. This follows immediately by applying the Pythagorean theorem to

$$f - \sum_{|n|\leq N} c_n e_n = f - S_N(f) + \sum_{|n|\leq N} b_n e_n,$$

where $b_n = a_n - c_n$.

This lemma has a clear geometric interpretation. It says that the trigonometric polynomial of degree at most N which is closest to f in the norm $\|\cdot\|$ is the partial sum $S_N(f)$. This geometric property of the partial sums is depicted in Figure 1, where the orthogonal projection of f in the plane spanned by $\{e_{-N}, \ldots, e_0, \ldots, e_N\}$ is simply $S_N(f)$.

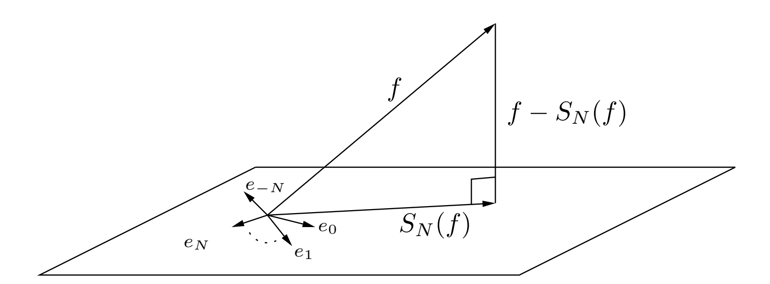


Figure 1. The best approximation lemma

We can now give the proof that $\|S_N(f) - f\| \to 0$ using the best approximation lemma, as well as the important fact that trigonometric polynomials are dense in the space of continuous functions on the circle.

Suppose that f is continuous on the circle. Then, given $\epsilon > 0$, there exists (by Corollary 5.4 in Chapter 2) a trigonometric polynomial P, say of degree M, such that

$$|f(\theta) - P(\theta)| < \epsilon \quad \text{for all } \theta.$$

In particular, taking squares and integrating this inequality yields $\|f - P\| < \epsilon$, and by the best approximation lemma we conclude that

$$\|f - S_N(f)\| < \epsilon \quad \text{whenever } N \geq M.$$

This proves Theorem 1.1 when f is continuous.

If f is merely integrable, we can no longer approximate f uniformly by trigonometric polynomials. Instead, we apply the approximation Lemma 3.2 in Chapter 2 and choose a continuous function g on the circle which satisfies

$$\sup_{\theta \in [0,2\pi]} |g(\theta)| \leq \sup_{\theta \in [0,2\pi]} |f(\theta)| = B,$$

and

$$\int_0^{2\pi} |f(\theta) - g(\theta)|\, d\theta < \epsilon^2.$$

Then we get

$$\begin{aligned}
\|f - g\|^2 &= \frac{1}{2\pi} \int_0^{2\pi} |f(\theta) - g(\theta)|^2\, d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} |f(\theta) - g(\theta)|\, |f(\theta) - g(\theta)|\, d\theta \\
&\leq \frac{2B}{2\pi} \int_0^{2\pi} |f(\theta) - g(\theta)|\, d\theta \\
&\leq C\epsilon^2.
\end{aligned}$$

Now we may approximate g by a trigonometric polynomial P so that $\|g - P\| < \epsilon$. Then $\|f - P\| < C'\epsilon$, and we may again conclude by applying the best approximation lemma. This completes the proof that the partial sums of the Fourier series of f converge to f in the mean square norm $\| \cdot \|$.

Note that this result and the relation (3) imply that if a_n is the nth Fourier coefficient of an integrable function f, then the series $\sum_{n=-\infty}^{\infty} |a_n|^2$ converges, and in fact we have **Parseval's identity**

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \|f\|^2.$$

This identity provides an important connection between the norms in the two vector spaces $\ell^2(\mathbb{Z})$ and $\mathcal{R}$.

We now summarize the results of this section.

Theorem 1.3 *Let f be an integrable function on the circle with $f \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$. Then we have:*

(i) *Mean-square convergence of the Fourier series*

$$\frac{1}{2\pi}\int_0^{2\pi} |f(\theta) - S_N(f)(\theta)|^2 \, d\theta \to 0 \quad \text{as } N \to \infty.$$

(ii) *Parseval's identity*

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|^2 \, d\theta.$$

Remark 1. If $\{e_n\}$ is *any* orthonormal family of functions on the circle, and $a_n = (f, e_n)$, then we may deduce from the relation (3) that

$$\sum_{n=-\infty}^{\infty} |a_n|^2 \leq \|f\|^2.$$

This is known as **Bessel's inequality**. Equality holds (as in Parseval's identity) precisely when the family $\{e_n\}$ is also a "basis," in the sense that $\|\sum_{|n|\leq N} a_n e_n - f\| \to 0$ as $N \to \infty$.

Remark 2. We may associate to every integrable function the sequence $\{a_n\}$ formed by its Fourier coefficients. Parseval's identity guarantees that $\{a_n\} \in \ell^2(\mathbb{Z})$. Since $\ell^2(\mathbb{Z})$ is a Hilbert space, the failure of $\mathcal{R}$ to be complete, discussed earlier, may be understood as follows: there exist sequences $\{a_n\}_{n\in\mathbb{Z}}$ such that $\sum_{n\in\mathbb{Z}} |a_n|^2 < \infty$, yet no Riemann integrable function F has n^{th} Fourier coefficient equal to a_n for all n. An example is given in Exercise 6.

Since the terms of a converging series tend to 0, we deduce from Parseval's identity or Bessel's inequality the following result.

Theorem 1.4 (Riemann-Lebesgue lemma) *If f is integrable on the circle, then $\hat{f}(n) \to 0$ as $|n| \to \infty$.*

An equivalent reformulation of this proposition is that if f is integrable on $[0, 2\pi]$, then

$$\int_0^{2\pi} f(\theta)\sin(N\theta)\, d\theta \to 0 \quad \text{as } N \to \infty$$

and

$$\int_0^{2\pi} f(\theta)\cos(N\theta)\,d\theta \to 0 \qquad \text{as } N \to \infty.$$

To conclude this section, we give a more general version of the Parseval identity which we will use in the next chapter.

Lemma 1.5 *Suppose F and G are integrable on the circle with*

$$F \sim \sum a_n e^{in\theta} \qquad \text{and} \qquad G \sim \sum b_n e^{in\theta}.$$

Then

$$\frac{1}{2\pi}\int_0^{2\pi} F(\theta)\overline{G(\theta)}\,d\theta = \sum_{n=-\infty}^{\infty} a_n\overline{b_n}.$$

Recall from the discussion in Example 1 that the series $\sum_{n=-\infty}^{\infty} a_n\overline{b_n}$ converges absolutely.

Proof. The proof follows from Parseval's identity and the fact that

$$(F,G) = \frac{1}{4}\left[\|F+G\|^2 - \|F-G\|^2 + i\left(\|F+iG\|^2 - \|F-iG\|^2\right)\right]$$

which holds in every Hermitian inner product space. The verification of this fact is left to the reader.

2 Return to pointwise convergence

The mean-square convergence theorem does not provide further insight into the problem of pointwise convergence. Indeed, Theorem 1.1 by itself does not guarantee that the Fourier series converges for any θ. Exercise 3 helps to explain this statement. However, if a function is differentiable at a point θ_0, then its Fourier series converges at θ_0. After proving this result, we give an example of a continuous function with diverging Fourier series at one point. These phenomena are indicative of the intricate nature of the problem of pointwise convergence in the theory of Fourier series.

2.1 A local result

Theorem 2.1 *Let f be an integrable function on the circle which is differentiable at a point θ_0. Then $S_N(f)(\theta_0) \to f(\theta_0)$ as N tends to infinity.*

Proof. Define

$$F(t) = \begin{cases} \dfrac{f(\theta_0 - t) - f(\theta_0)}{t} & \text{if } t \neq 0 \text{ and } |t| < \pi \\ -f'(\theta_0) & \text{if } t = 0. \end{cases}$$

First, F is bounded near 0 since f is differentiable there. Second, for all small δ the function F is integrable on $[-\pi, -\delta] \cup [\delta, \pi]$ because f has this property and $|t| > \delta$ there. As a consequence of Proposition 1.4 in the appendix, the function F is integrable on all of $[-\pi, \pi]$. We know that $S_N(f)(\theta_0) = (f * D_N)(\theta_0)$, where D_N is the Dirichlet kernel. Since $\frac{1}{2\pi}\int D_N = 1$, we find that

$$\begin{aligned} S_N(f)(\theta_0) - f(\theta_0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta_0 - t)D_N(t)\,dt - f(\theta_0) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} [f(\theta_0 - t) - f(\theta_0)]D_N(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} F(t)tD_N(t)\,dt. \end{aligned}$$

We recall that

$$tD_N(t) = \frac{t}{\sin(t/2)}\sin((N + 1/2)t),$$

where the quotient $\frac{t}{\sin(t/2)}$ is continuous in the interval $[-\pi, \pi]$. Since we can write

$$\sin((N + 1/2)t) = \sin(Nt)\cos(t/2) + \cos(Nt)\sin(t/2),$$

we can apply the Riemann-Lebesgue lemma to the Riemann integrable functions $F(t)t\cos(t/2)/\sin(t/2)$ and $F(t)t$ to finish the proof of the theorem.

Observe that the conclusion of the theorem still holds if we only assume that f satisfies a **Lipschitz condition** at θ_0; that is,

$$|f(\theta) - f(\theta_0)| \leq M|\theta - \theta_0|$$

for some $M \geq 0$ and all θ. This is the same as saying that f satisfies a Hölder condition of order $\alpha = 1$.

A striking consequence of this theorem is the localization principle of Riemann. This result states that the convergence of $S_N(f)(\theta_0)$ depends only on the behavior of f near θ_0. This is not clear at first, since forming the Fourier series requires integrating f over the *whole* circle.

Theorem 2.2 *Suppose f and g are two integrable functions defined on the circle, and for some θ_0 there exists an open interval I containing θ_0 such that*

$$f(\theta) = g(\theta) \quad \text{for all } \theta \in I.$$

Then $S_N(f)(\theta_0) - S_N(g)(\theta_0) \to 0$ as N tends to infinity.

Proof. The function $f - g$ is 0 in I, so it is differentiable at θ_0, and we may apply the previous theorem to conclude the proof.

2.2 A continuous function with diverging Fourier series

We now turn our attention to an example of a continuous periodic function whose Fourier series diverges at a point. Thus, Theorem 2.1 fails if the differentiability assumption is replaced by the weaker assumption of continuity. Our counter-example shows that this hypothesis which had appeared plausible, is in fact false; moreover, its construction also illuminates an important principle of the theory.

The principle that is involved here will be referred to as "symmetry-breaking."[1] The symmetry that we have in mind is the symmetry between the frequencies $e^{in\theta}$ and $e^{-in\theta}$ which appear in the Fourier expansion of a function. For example, the partial sum operator S_N is defined in a way that reflects this symmetry. Also, the Dirichlet, Fejèr, and Poisson kernels are symmetric in this sense. When we break the symmetry, that is, when we split the Fourier series $\sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ into the two pieces $\sum_{n\geq 0} a_n e^{in\theta}$ and $\sum_{n<0} a_n e^{in\theta}$, we introduce new and far-reaching phenomena.

We give a simple example. Start with the sawtooth function f which is odd in θ and which equals $i(\pi - \theta)$ when $0 < \theta < \pi$. Then, by Exercise 8 in Chapter 2, we know that

$$f(\theta) \sim \sum_{n\neq 0} \frac{e^{in\theta}}{n}. \tag{4}$$

Consider now the result of breaking the symmetry and the resulting series

$$\sum_{n=-\infty}^{n=-1} \frac{e^{in\theta}}{n}.$$

Then, unlike (4), the above is no longer the Fourier series of a Riemann integrable function. Indeed, suppose it were the Fourier series of an

[1] We have borrowed this terminology from physics, where it is used in a very different context.

integrable function, say $\tilde{f}$, where in particular $\tilde{f}$ is bounded. Using the Abel means, we then have

$$|A_r(\tilde{f})(0)| = \sum_{n=1}^{\infty} \frac{r^n}{n},$$

which tends to infinity as r tends to 1, because $\sum 1/n$ diverges. This gives the desired contradiction since

$$|A_r(\tilde{f})(0)| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |\tilde{f}(\theta)| P_r(\theta)\, d\theta \leq \sup_{\theta} \ |\tilde{f}(\theta)|,$$

where $P_r(\theta)$ denotes the Poisson kernel discussed in the previous chapter.

The sawtooth function is the object from which we will fashion our counter-example. We proceed as follows. For each $N \geq 1$ we define the following two functions on $[-\pi, \pi]$,

$$f_N(\theta) = \sum_{1\leq |n| \leq N} \frac{e^{in\theta}}{n} \quad \text{and} \quad \tilde{f}_N(\theta) = \sum_{-N\leq n\leq -1} \frac{e^{in\theta}}{n}.$$

We contend that:

(i) $|\tilde{f}_N(0)| \geq c \log N$.

(ii) $f_N(\theta)$ is uniformly bounded in N and θ.

The first statement is a consequence of the fact that $\sum_{n=1}^{N} 1/n \geq \log N$, which is easily established (see also Figure 2):

$$\sum_{n=1}^{N} \frac{1}{n} \geq \sum_{n=1}^{N-1} \int_{n}^{n+1} \frac{dx}{x} = \int_{1}^{N} \frac{dx}{x} = \log N.$$

To prove (ii), we argue in the same spirit as in the proof of Tauber's theorem, which says that if the series $\sum c_n$ is Abel summable to s and $c_n = o(1/n)$, then $\sum c_n$ actually converges to s (see Exercise 14 in Chapter 2). In fact, the proof of Tauber's theorem is quite similar to that of the lemma below.

Lemma 2.3 *Suppose that the Abel means $A_r = \sum_{n=1}^{\infty} r^n c_n$ of the series $\sum_{n=1}^{\infty} c_n$ are bounded as r tends to 1 (with $r < 1$). If $c_n = O(1/n)$, then the partial sums $S_N = \sum_{n=1}^{N} c_n$ are bounded.*

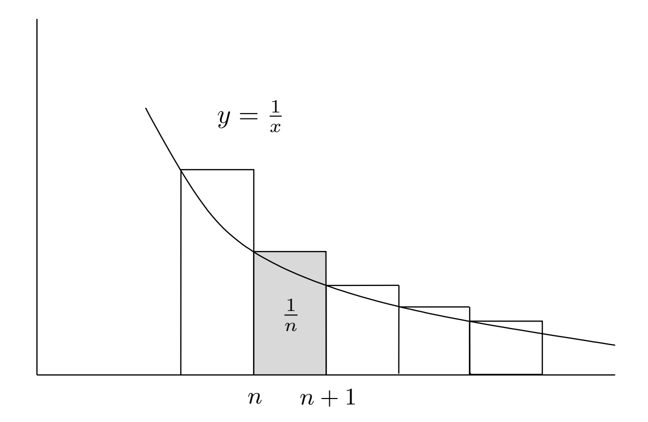


Figure 2. Comparing a sum with an integral

Proof. Let $r = 1 - 1/N$ and choose M so that $n|c_n| \leq M$. We estimate the difference

$$S_N - A_r = \sum_{n=1}^{N}(c_n - r^n c_n) - \sum_{n=N+1}^{\infty} r^n c_n$$

as follows:

$$\begin{aligned}
|S_N - A_r| &\leq \sum_{n=1}^{N} |c_n|(1 - r^n) + \sum_{n=N+1}^{\infty} r^n |c_n| \\
&\leq M \sum_{n=1}^{N}(1 - r) + \frac{M}{N} \sum_{n=N+1}^{\infty} r^n \\
&\leq MN(1 - r) + \frac{M}{N} \frac{1}{1 - r} \\
&= 2M,
\end{aligned}$$

where we have used the simple observation that

$$1 - r^n = (1 - r)(1 + r + \cdots + r^{n-1}) \leq n(1 - r).$$

So we see that if M satisfies both $|A_r| \leq M$ and $n|c_n| \leq M$, then $|S_N| \leq 3M$.

We apply the lemma to the series

$$\sum_{n \neq 0} \frac{e^{in\theta}}{n},$$

which is the Fourier series of the sawtooth function f used above. Here $c_n = e^{in\theta}/n + e^{-in\theta}/(-n)$ for $n \neq 0$, so clearly $c_n = O(1/|n|)$. Finally, the Abel means of this series are $A_r(f)(\theta) = (f * P_r)(\theta)$. But f is bounded and P_r is a good kernel, so $S_N(f)(\theta)$ is uniformly bounded in N and θ, as was to be shown.

We now come to the heart of the matter. Notice that f_N and $\tilde{f}_N$ are trigonometric polynomials of degree N (that is, they have non-zero Fourier coefficients only when $|n| \leq N$). From these, we form trigonometric polynomials P_N and $\tilde{P}_N$, now of degrees $3N$ and $2N-1$, by displacing the frequencies of f_N and $\tilde{f}_N$ by $2N$ units. In other words, we define $P_N(\theta) = e^{i(2N)\theta} f_N(\theta)$ and $\tilde{P}_N(\theta) = e^{i(2N)\theta} \tilde{f}_N(\theta)$. So while f_N has non-vanishing Fourier coefficients when $0 < |n| \leq N$, now the coefficients of P_N are non-vanishing for $N \leq n \leq 3N$, $n \neq 2N$. Moreover, while $n = 0$ is the center of symmetry of f_N, now $n = 2N$ is the center of symmetry of P_N. We next consider the partial sums S_M.

Lemma 2.4

$$S_M(P_N) = \begin{cases} P_N & \textit{if } M \geq 3N, \\ \tilde{P}_N & \textit{if } M = 2N, \\ 0 & \textit{if } M < N. \end{cases}$$

This is clear from what has been said above and from Figure 3.

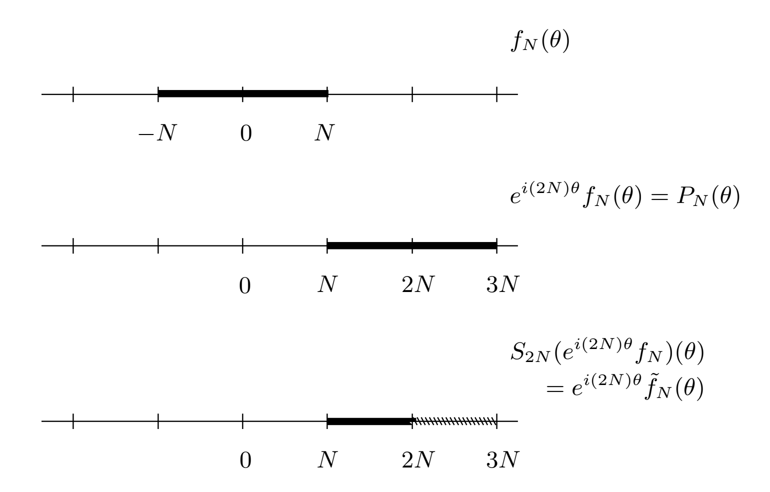


Figure 3. Breaking symmetry in Lemma 2.4

The effect is that when $M = 2N$, the operator S_M breaks the symmetry of P_N, but in the other cases covered in the lemma, the action of S_M

is relatively benign, since then the outcome is either P_N or 0.

Finally, we need to find a convergent series of positive terms $\sum \alpha_k$ and a sequence of integers $\{N_k\}$ which increases rapidly enough so that:

(i) $N_{k+1} > 3N_k$,

(ii) $\alpha_k \log N_k \to \infty$ as $k \to \infty$.

We choose (for example) $\alpha_k = 1/k^2$ and $N_k = 3^{2^k}$ which are easily seen to satisfy the above criteria.

Finally, we can write down our desired function. It is

$$f(\theta) = \sum_{k=1}^{\infty} \alpha_k P_{N_k}(\theta).$$

Due to the uniform boundedness of the P_N (recall that $|P_N(\theta)| = |f_N(\theta)|$), the series above converges uniformly to a continuous periodic function. However, by our lemma we get

$$|S_{2N_m}(f)(0)| \geq c\alpha_m \log N_m + O(1) \to \infty \qquad \text{as } m \to \infty.$$

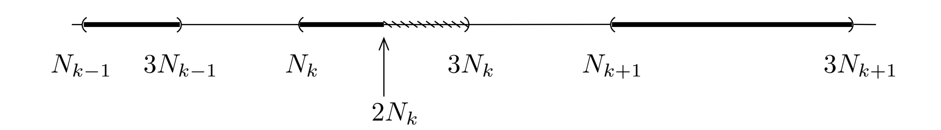


Figure 4. Symmetry broken in the middle interval $(N_k, 3N_k)$

Indeed, the terms that correspond to N_k with $k < m$ or $k > m$ contribute $O(1)$ or 0, respectively (because the P_N's are uniformly bounded), while the term that corresponds to N_m is in absolute value greater than $c\alpha_m \log N_m$ because $|\tilde{P}_N(\theta)| = |\tilde{f}_N(\theta)| \geq c \log N$. So the partial sums of the Fourier series of f at 0 are not bounded, and we are done since this proves the divergence of the Fourier series of f at $\theta = 0$. To produce a function whose series diverges at any other preassigned $\theta = \theta_0$, it suffices to consider the function $f(\theta - \theta_0)$.

3 Exercises

1. Show that the first two examples of inner product spaces, namely $\mathbb{R}^d$ and $\mathbb{C}^d$, are complete.

[Hint: Every Cauchy sequence in $\mathbb{R}$ has a limit.]

2. Prove that the vector space $\ell^2(\mathbb{Z})$ is complete.

[Hint: Suppose $A_k = \{a_{k,n}\}_{n\in\mathbb{Z}}$ with $k = 1, 2, \ldots$ is a Cauchy sequence. Show that for each n, $\{a_{k,n}\}_{k=1}^{\infty}$ is a Cauchy sequence of complex numbers, therefore it converges to a limit, say b_n. By taking partial sums of $\|A_k - A_{k'}\|$ and letting $k' \to \infty$, show that $\|A_k - B\| \to 0$ as $k \to \infty$, where $B = (\ldots, b_{-1}, b_0, b_1, \ldots)$. Finally, prove that $B \in \ell^2(\mathbb{Z})$.]

3. Construct a sequence of integrable functions $\{f_k\}$ on $[0, 2\pi]$ such that

$$\lim_{k\to\infty} \frac{1}{2\pi}\int_0^{2\pi} |f_k(\theta)|^2\, d\theta = 0$$

but $\lim_{k\to\infty} f_k(\theta)$ fails to exist for any θ.

[Hint: Choose a sequence of intervals $I_k \subset [0, 2\pi]$ whose lengths tend to 0, and so that each point belongs to infinitely many of them; then let $f_k = \chi_{I_k}$.]

4. Recall the vector space $\mathcal{R}$ of integrable functions, with its inner product and norm

$$\|f\| = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\, dx\right)^{1/2}.$$

(a) Show that there exist non-zero integrable functions f for which $\|f\| = 0$.

(b) However, show that if $f \in \mathcal{R}$ with $\|f\| = 0$, then $f(x) = 0$ whenever f is continuous at x.

(c) Conversely, show that if $f \in \mathcal{R}$ vanishes at all of its points of continuity, then $\|f\| = 0$.

5. Let

$$f(\theta) = \begin{cases} 0 & \text{for } \theta = 0 \\ \log(1/\theta) & \text{for } 0 < \theta \leq 2\pi, \end{cases}$$

and define a sequence of functions in $\mathcal{R}$ by

$$f_n(\theta) = \begin{cases} 0 & \text{for } 0 \leq \theta \leq 1/n \\ f(\theta) & \text{for } 1/n < \theta \leq 2\pi. \end{cases}$$

Prove that $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{R}$. However, f does not belong to $\mathcal{R}$.

[Hint: Show that $\int_a^b (\log\theta)^2\,d\theta \to 0$ if $0 < a < b$ and $b \to 0$, by using the fact that the derivative of $\theta(\log\theta)^2 - 2\theta\log\theta + 2\theta$ is equal to $(\log\theta)^2$.]

6. Consider the sequence $\{a_k\}_{k=-\infty}^{\infty}$ defined by

$$a_k = \begin{cases} 1/k & \text{if } k \geq 1 \\ 0 & \text{if } k \leq 0. \end{cases}$$

Note that $\{a_k\} \in \ell^2(\mathbb{Z})$, but that no Riemann integrable function has k^{th} Fourier coefficient equal to a_k for all k.

7. Show that the trigonometric series

$$\sum_{n\geq 2} \frac{1}{\log n} \sin nx$$

converges for every x, yet it is not the Fourier series of a Riemann integrable function.

The same is true for $\sum \frac{\sin nx}{n^\alpha}$ for $0 < \alpha < 1$, but the case $1/2 < \alpha < 1$ is more difficult. See Problem 1.

8. Exercise 6 in Chapter 2 dealt with the sums

$$\sum_{n \text{ odd } \geq 1} \frac{1}{n^2} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Similar sums can be derived using the methods of this chapter.

(a) Let f be the function defined on $[-\pi, \pi]$ by $f(\theta) = |\theta|$. Use Parseval's identity to find the sums of the following two series:

$$\sum_{n=0}^{\infty} \frac{1}{(2n+1)^4} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^4}.$$

In fact, they are $\pi^4/96$ and $\pi^4/90$, respectively.

(b) Consider the 2π-periodic odd function defined on $[0, \pi]$ by $f(\theta) = \theta(\pi - \theta)$. Show that

$$\sum_{n=0}^{\infty} \frac{1}{(2n+1)^6} = \frac{\pi^6}{960} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^6}{945}.$$

Remark. The general expression when k is even for $\sum_{n=1}^{\infty} 1/n^k$ in terms of π^k is given in Problem 4. However, finding a formula for the sum $\sum_{n=1}^{\infty} 1/n^3$, or more generally $\sum_{n=1}^{\infty} 1/n^k$ with k odd, is a famous unresolved question.

9. Show that for α not an integer, the Fourier series of

$$\frac{\pi}{\sin \pi\alpha} e^{i(\pi - x)\alpha}$$

on $[0, 2\pi]$ is given by

$$\sum_{n=-\infty}^{\infty} \frac{e^{inx}}{n+\alpha}.$$

Apply Parseval's formula to show that

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\alpha)^2} = \frac{\pi^2}{(\sin \pi\alpha)^2}.$$

10. Consider the example of a vibrating string which we analyzed in Chapter 1. The displacement $u(x,t)$ of the string at time t satisfies the wave equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad c^2 = \tau/\rho.$$

The string is subject to the initial conditions

$$u(x,0) = f(x) \qquad \text{and} \qquad \frac{\partial u}{\partial t}(x,0) = g(x),$$

where we assume that $f \in C^1$ and g is continuous. We define the total **energy** of the string by

$$E(t) = \frac{1}{2}\rho \int_0^L \left(\frac{\partial u}{\partial t}\right)^2 dx + \frac{1}{2}\tau \int_0^L \left(\frac{\partial u}{\partial x}\right)^2 dx.$$

The first term corresponds to the "kinetic energy" of the string (in analogy with $(1/2)mv^2$, the kinetic energy of a particle of mass m and velocity v), and the second term corresponds to its "potential energy."

Show that the total energy of the string is conserved, in the sense that $E(t)$ is constant. Therefore,

$$E(t) = E(0) = \frac{1}{2}\rho \int_0^L g(x)^2\, dx + \frac{1}{2}\tau \int_0^L f'(x)^2\, dx.$$

11. The inequalities of Wirtinger and Poincaré establish a relationship between the norm of a function and that of its derivative.

(a) If f is T-periodic, continuous, and piecewise C^1 with $\int_0^T f(t)\,dt = 0$, show that

$$\int_0^T |f(t)|^2\,dt \le \frac{T^2}{4\pi^2}\int_0^T |f'(t)|^2\,dt,$$

with equality if and only if $f(t) = A\sin(2\pi t/T) + B\cos(2\pi t/T)$. [Hint: Apply Parseval's identity.]

(b) If f is as above and g is just C^1 and T-periodic, prove that

$$\left|\int_0^T \overline{f(t)}g(t)\,dt\right|^2 \le \frac{T^2}{4\pi^2}\int_0^T |f(t)|^2\,dt \int_0^T |g'(t)|^2\,dt.$$

(c) For any compact interval $[a,b]$ and any continuously differentiable function f with $f(a) = f(b) = 0$, show that

$$\int_a^b |f(t)|^2\,dt \le \frac{(b-a)^2}{\pi^2}\int_a^b |f'(t)|^2\,dt.$$

Discuss the case of equality, and prove that the constant $(b-a)^2/\pi^2$ cannot be improved. [Hint: Extend f to be odd with respect to a and periodic of period $T = 2(b-a)$ so that its integral over an interval of length T is 0. Apply part a) to get the inequality, and conclude that equality holds if and only if $f(t) = A\sin(\pi\frac{t-a}{b-a})$].

12. Prove that $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$.

[Hint: Start with the fact that the integral of $D_N(\theta)$ equals 2π, and note that the difference $(1/\sin(\theta/2)) - 2/\theta$ is continuous on $[-\pi,\pi]$. Apply the Riemann-Lebesgue lemma.]

13. Suppose that f is periodic and of class C^k. Show that

$$\hat{f}(n) = o(1/|n|^k),$$

that is, $|n|^k\hat{f}(n)$ goes to 0 as $|n|\to\infty$. This is an improvement over Exercise 10 in Chapter 2.

[Hint: Use the Riemann-Lebesgue lemma.]

14. Prove that the Fourier series of a continuously differentiable function f on the circle is absolutely convergent.

[Hint: Use the Cauchy-Schwarz inequality and Parseval's identity for f'.]

15. Let f be 2π-periodic and Riemann integrable on $[-\pi,\pi]$.

(a) Show that

$$\hat{f}(n) = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+\pi/n)e^{-inx}\,dx$$

hence

$$\hat{f}(n) = \frac{1}{4\pi}\int_{-\pi}^{\pi} [f(x) - f(x+\pi/n)]e^{-inx}\,dx.$$

(b) Now assume that f satisfies a Hölder condition of order α, namely

$$|f(x+h) - f(x)| \leq C|h|^{\alpha}$$

for some $0 < \alpha \leq 1$, some $C > 0$, and all x, h. Use part a) to show that

$$\hat{f}(n) = O(1/|n|^{\alpha}).$$

(c) Prove that the above result cannot be improved by showing that the function

$$f(x) = \sum_{k=0}^{\infty} 2^{-k\alpha} e^{i2^k x},$$

where $0 < \alpha < 1$, satisfies

$$|f(x+h) - f(x)| \leq C|h|^{\alpha},$$

and $\hat{f}(N) = 1/N^{\alpha}$ whenever $N = 2^k$.

[Hint: For (c), break up the sum as follows $f(x+h) - f(x) = \sum_{2^k \leq 1/|h|} + \sum_{2^k > 1/|h|}$. To estimate the first sum use the fact that $|1 - e^{i\theta}| \leq |\theta|$ whenever θ is small. To estimate the second sum, use the obvious inequality $|e^{ix} - e^{iy}| \leq 2$.]

16. Let f be a 2π-periodic function which satisfies a Lipschitz condition with constant K; that is,

$$|f(x) - f(y)| \leq K|x - y| \qquad \text{for all } x, y.$$

This is simply the Hölder condition with $\alpha = 1$, so by the previous exercise, we see that $\hat{f}(n) = O(1/|n|)$. Since the harmonic series $\sum 1/n$ diverges, we cannot say anything (yet) about the absolute convergence of the Fourier series of f. The outline below actually proves that the Fourier series of f converges absolutely and uniformly.

(a) For every positive h we define $g_h(x) = f(x+h) - f(x-h)$. Prove that

$$\frac{1}{2\pi}\int_0^{2\pi} |g_h(x)|^2\, dx = \sum_{n=-\infty}^{\infty} 4|\sin nh|^2 |\hat{f}(n)|^2,$$

and show that

$$\sum_{n=-\infty}^{\infty} |\sin nh|^2 |\hat{f}(n)|^2 \leq K^2 h^2.$$

(b) Let p be a positive integer. By choosing $h = \pi/2^{p+1}$, show that

$$\sum_{2^{p-1}<|n|\leq 2^p} |\hat{f}(n)|^2 \leq \frac{K^2\pi^2}{2^{2p+1}}.$$

(c) Estimate $\sum_{2^{p-1}<|n|\leq 2^p} |\hat{f}(n)|$, and conclude that the Fourier series of f converges absolutely, hence uniformly. [Hint: Use the Cauchy-Schwarz inequality to estimate the sum.]

(d) In fact, modify the argument slightly to prove Bernstein's theorem: If f satisfies a Hölder condition of order $\alpha > 1/2$, then the Fourier series of f converges absolutely.

17. If f is a bounded monotonic function on $[-\pi, \pi]$, then

$$\hat{f}(n) = O(1/|n|).$$

[Hint: One may assume that f is increasing, and say $|f| \leq M$. First check that the Fourier coefficients of the characteristic function of $[a, b]$ satisfy $O(1/|n|)$. Now show that a sum of the form

$$\sum_{k=1}^{N} \alpha_k \chi_{[a_k, a_{k+1}]}(x)$$

with $-\pi = a_1 < a_2 < \cdots < a_N < a_{N+1} = \pi$ and $-M \leq \alpha_1 \leq \cdots \leq \alpha_N \leq M$ has Fourier coefficients that are $O(1/|n|)$ uniformly in N. Summing by parts one gets a telescopic sum $\sum(\alpha_{k+1} - \alpha_k)$ which can be bounded by $2M$. Now approximate f by functions of the above type.]

18. Here are a few things we have learned about the decay of Fourier coefficients:

(a) if f is of class C^k, then $\hat{f}(n) = o(1/|n|^k)$;

(b) if f is Lipschitz, then $\hat{f}(n) = O(1/|n|)$;

(c) if f is monotonic, then $\hat{f}(n) = O(1/|n|)$;

(d) if f is satisfies a Hölder condition with exponent α where $0 < \alpha < 1$, then $\hat{f}(n) = O(1/|n|^\alpha)$;

(e) if f is merely Riemann integrable, then $\sum |\hat{f}(n)|^2 < \infty$ and therefore $\hat{f}(n) = o(1)$.

Nevertheless, show that the Fourier coefficients of a continuous function can tend to 0 arbitrarily slowly by proving that for every sequence of nonnegative real numbers $\{\epsilon_n\}$ converging to 0, there exists a continuous function f such that $|\hat{f}(n)| \geq \epsilon_n$ for infinitely many values of n.

[Hint: Choose a subsequence $\{\epsilon_{n_k}\}$ so that $\sum_k \epsilon_{n_k} < \infty$.]

19. Give another proof that the sum $\sum_{0<|n|\leq N} e^{inx}/n$ is uniformly bounded in N and $x \in [-\pi, \pi]$ by using the fact that

$$\frac{1}{2i} \sum_{0<|n|\leq N} \frac{e^{inx}}{n} = \sum_{n=1}^{N} \frac{\sin nx}{n} = \frac{1}{2} \int_0^x (D_N(t) - 1)\, dt,$$

where D_N is the Dirichlet kernel. Now use the fact that $\int_0^\infty \frac{\sin t}{t}\, dt < \infty$ which was proved in Exercise 12.

20. Let $f(x)$ denote the sawtooth function defined by $f(x) = (\pi - x)/2$ on the interval $(0, 2\pi)$ with $f(0) = 0$ and extended by periodicity to all of $\mathbb{R}$. The Fourier series of f is

$$f(x) \sim \frac{1}{2i} \sum_{|n|\neq 0} \frac{e^{inx}}{n} = \sum_{n=1}^{\infty} \frac{\sin nx}{n},$$

and f has a jump discontinuity at the origin with

$$f(0^+) = \frac{\pi}{2}, \qquad f(0^-) = -\frac{\pi}{2}, \qquad \text{and hence} \qquad f(0^+) - f(0^-) = \pi.$$

Show that

$$\max_{0<x\leq \pi/N} S_N(f)(x) - \frac{\pi}{2} = \int_0^\pi \frac{\sin t}{t}\, dt - \frac{\pi}{2},$$

which is roughly 9% of the jump π. This result is a manifestation of Gibbs's phenomenon which states that near a jump discontinuity, the Fourier series of a function overshoots (or undershoots) it by approximately 9% of the jump.

[Hint: Use the expression for $S_N(f)$ given in Exercise 19.]

4 Problems

1. For each $0 < \alpha < 1$ the series

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n^\alpha}$$

converges for every x but is not the Fourier series of a Riemann integrable function.

(a) If the **conjugate Dirichlet kernel** is defined by

$$\tilde{D}_N(x) = \sum_{|n|\leq N} \text{sign}(x)\, e^{inx} \qquad \text{where } \text{sign}(x) = \begin{cases} 1 & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -1 & \text{if } n < 0, \end{cases}$$

then show that

$$\tilde{D}_N(x) = \frac{\cos(x/2) - \cos((N+1/2)x)}{\sin(x/2)},$$

and

$$\int_{-\pi}^{\pi} |\tilde{D}_N(x)|\, dx \leq c \log N.$$

(b) As a result, if f is Riemann integrable, then

$$(f * \tilde{D}_N)(0) = O(\log N).$$

(c) In the present case, this leads to

$$\sum_{n=1}^{N} \frac{1}{n^\alpha} = O(\log N),$$

which is a contradiction.

2. An important fact we have proved is that the family $\{e^{inx}\}_{n\in\mathbb{Z}}$ is orthonormal in $\mathcal{R}$ and it is also complete, in the sense that the Fourier series of f converges to f in the norm. In this exercise, we consider another family possessing these same properties.

On $[-1, 1]$ define

$$L_n(x) = \frac{d^n}{dx^n}(x^2-1)^n, \qquad n = 0, 1, 2, \ldots.$$

Then L_n is a polynomial of degree n which is called the nth **Legendre polynomial**.

(a) Show that if f is indefinitely differentiable on $[-1, 1]$, then

$$\int_{-1}^{1} L_n(x) f(x)\, dx = (-1)^n \int_{-1}^{1} (x^2 - 1)^n f^{(n)}(x)\, dx.$$

In particular, show that L_n is orthogonal to x^m whenever $m < n$. Hence $\{L_n\}_{n=0}^{\infty}$ is an orthogonal family.

(b) Show that

$$\|L_n\|^2 = \int_{-1}^{1} |L_n(x)|^2\, dx = \frac{(n!)^2 2^{2n+1}}{2n+1}.$$

[Hint: First, note that $\|L_n\|^2 = (-1)^n (2n)! \int_{-1}^{1} (x^2 - 1)^n\, dx$. Write $(x^2 - 1)^n = (x - 1)^n (x + 1)^n$ and integrate by parts n times to calculate this last integral.]

(c) Prove that any polynomial of degree n that is orthogonal to $1, x, x^2, \dots, x^{n-1}$ is a constant multiple of L_n.

(d) Let $\mathcal{L}_n = L_n / \|L_n\|$, which are the normalized Legendre polynomials. Prove that $\{\mathcal{L}_n\}$ is the family obtained by applying the "Gram-Schmidt process" to $\{1, x, \dots, x^n, \dots\}$, and conclude that every Riemann integrable function f on $[-1, 1]$ has a **Legendre expansion**

$$\sum_{n=0}^{\infty} \langle f, \mathcal{L}_n \rangle \mathcal{L}_n$$

which converges to f in the mean-square sense.

3. Let α be a complex number not equal to an integer.

(a) Calculate the Fourier series of the 2π-periodic function defined on $[-\pi, \pi]$ by $f(x) = \cos(\alpha x)$.

(b) Prove the following formulas due to Euler:

$$\sum_{n=1}^{\infty} \frac{1}{n^2 - \alpha^2} = \frac{1}{2\alpha^2} - \frac{\pi}{2\alpha \tan(\alpha \pi)}.$$

For all $u \in \mathbb{C} - \pi\mathbb{Z}$,

$$\cot u = \frac{1}{u} + 2 \sum_{n=1}^{\infty} \frac{u}{u^2 - n^2 \pi^2}.$$

(c) Show that for all $\alpha \in \mathbb{C} - \mathbb{Z}$ we have

$$\frac{\alpha\pi}{\sin(\alpha\pi)} = 1 + 2\alpha^2 \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2 - \alpha^2}.$$

(d) For all $0 < \alpha < 1$, show that

$$\int_0^{\infty} \frac{t^{\alpha-1}}{t+1}\, dt = \frac{\pi}{\sin(\alpha\pi)}.$$

[Hint: Split the integral as $\int_0^1 + \int_1^{\infty}$ and change variables $t = 1/u$ in the second integral. Now both integrals are of the form

$$\int_0^1 \frac{t^{\gamma-1}}{1+t}\, dt, \qquad 0 < \gamma < 1,$$

which one can show is equal to $\sum_{k=0}^{\infty} \frac{(-1)^k}{k+\gamma}$. Use part (c) to conclude the proof.]

4. In this problem, we find the formula for the sum of the series

$$\sum_{n=1}^{\infty} \frac{1}{n^k}$$

where k is any even integer. These sums are expressed in terms of the Bernoulli numbers; the related Bernoulli polynomials are discussed in the next problem.

Define the **Bernoulli numbers** B_n by the formula

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n.$$

(a) Show that $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, and $B_5 = 0$.

(b) Show that for $n \geq 1$ we have

$$B_n = -\frac{1}{n+1} \sum_{k=0}^{n-1} \binom{n+1}{k} B_k.$$

(c) By writing

$$\frac{z}{e^z - 1} = 1 - \frac{z}{2} + \sum_{n=2}^{\infty} \frac{B_n}{n!} z^n,$$

show that $B_n = 0$ if n is odd and > 1. Also prove that

$$z \cot z = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{2^{2n} B_{2n}}{(2n)!} z^{2n}.$$

(d) The **zeta function** is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad \text{for all } s > 1.$$

Deduce from the result in (c), and the expression for the cotangent function obtained in the previous problem, that

$$x \cot x = 1 - 2 \sum_{m=1}^{\infty} \frac{\zeta(2m)}{\pi^{2m}} x^{2m}.$$

(e) Conclude that

$$2\zeta(2m) = (-1)^{m+1} \frac{(2\pi)^{2m}}{(2m)!} B_{2m}.$$

5. Define the **Bernoulli polynomials** $B_n(x)$ by the formula

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n.$$

(a) The functions $B_n(x)$ are polynomials in x and

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k x^{n-k}.$$

Show that $B_0(x) = 1$, $B_1(x) = x - 1/2$, $B_2(x) = x^2 - x + 1/6$, and $B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x$.

(b) If $n \geq 1$, then

$$B_n(x+1) - B_n(x) = nx^{n-1},$$

and if $n \geq 2$, then

$$B_n(0) = B_n(1) = B_n.$$

(c) Define $S_m(n) = 1^m + 2^m + \cdots + (n-1)^m$. Show that

$$(m+1)S_m(n) = B_{m+1}(n) - B_{m+1}.$$

(d) Prove that the Bernoulli polynomials are the only polynomials that satisfy

(i) $B_0(x) = 1$,

(ii) $B_n'(x) = nB_{n-1}(x)$ for $n \geq 1$,

(iii) $\int_0^1 B_n(x)\,dx = 0$ for $n \geq 1$, and show that from (b) one obtains

$$\int_x^{x+1} B_n(t)\,dt = x^n.$$

(e) Calculate the Fourier series of $B_1(x)$ to conclude that for $0 < x < 1$ we have

$$B_1(x) = x - 1/2 = \frac{-1}{\pi}\sum_{k=1}^{\infty}\frac{\sin(2\pi kx)}{k}.$$

Integrate and conclude that

$$B_{2n}(x) = (-1)^{n+1}\frac{2(2n)!}{(2\pi)^{2n}}\sum_{k=1}^{\infty}\frac{\cos(2\pi kx)}{k^{2n}},$$

$$B_{2n+1}(x) = (-1)^{n+1}\frac{2(2n+1)!}{(2\pi)^{2n+1}}\sum_{k=1}^{\infty}\frac{\sin(2\pi kx)}{k^{2n+1}}.$$

Finally, show that for $0 < x < 1$,

$$B_n(x) = -\frac{n!}{(2\pi i)^n}\sum_{k\neq 0}\frac{e^{2\pi ikx}}{k^n}.$$

We observe that the Bernoulli polynomials are, up to normalization, successive integrals of the sawtooth function.

4 Some Applications of Fourier Series

> Fourier series and analogous expansions intervene very naturally in the general theory of curves and surfaces. In effect, this theory, conceived from the point of view of analysis, deals obviously with the study of arbitrary functions. I was thus led to use Fourier series in several questions of geometry, and I have obtained in this direction a number of results which will be presented in this work. One notes that my considerations form only a beginning of a principal series of researches, which would without doubt give many new results.
>
> *A. Hurwitz,* 1902

In the previous chapters we introduced some basic facts about Fourier analysis, motivated by problems that arose in physics. The motion of a string and the diffusion of heat were two instances that led naturally to the expansion of a function in terms of a Fourier series. We propose next to give the reader a flavor of the broader impact of Fourier analysis, and illustrate how these ideas reach out to other areas of mathematics. In particular, consider the following three problems:

I. Among all simple closed curves of length ℓ in the plane $\mathbb{R}^2$, which one encloses the largest area?

II. Given an irrational number γ, what can be said about the distribution of the fractional parts of the sequence of numbers $n\gamma$, for $n = 1, 2, 3, \ldots$?

III. Does there exist a continuous function that is nowhere differentiable?

The first problem is clearly geometric in nature, and at first sight, would seem to have little to do with Fourier series. The second question lies on the border between number theory and the study of dynamical systems, and gives us the simplest example of the idea of "ergodicity." The third problem, while analytic in nature, resisted many attempts before the

solution was finally discovered. It is remarkable that all three questions can be resolved quite simply and directly by the use of Fourier series.

In the last section of this chapter, we return to a problem that provided our initial motivation. We consider the time-dependent heat equation on the circle. Here our investigation will lead us to the important but enigmatic heat kernel for the circle. However, the mysteries surrounding its basic properties will not be fully understood until we can apply the Poisson summation formula, which we will do in the next chapter.

1 The isoperimetric inequality

Let Γ denote a closed curve in the plane which does not intersect itself. Also, let ℓ denote the length of Γ, and $\mathcal{A}$ the area of the bounded region in $\mathbb{R}^2$ enclosed by Γ. The problem now is to determine for a given ℓ the curve Γ which maximizes $\mathcal{A}$ (if any such curve exists).

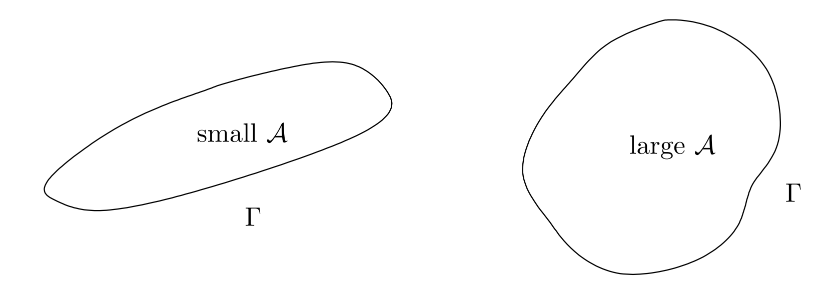


Figure 1. The isoperimetric problem

A little experimentation and reflection suggests that the solution should be a circle. This conclusion can be reached by the following heuristic considerations. The curve can be thought of as a closed piece of string lying flat on a table. If the region enclosed by the string is not convex (for example), one can deform part of the string and increase the area enclosed by it. Also, playing with some simple examples, one can convince oneself that the "flatter" the curve is in some portion, the less efficient it is in enclosing area. Therefore we want to maximize the "roundness" of the curve at each point.

Although the circle is the correct guess, making the above ideas precise is a difficult matter.

The key idea in the solution we give to the isoperimetric problem consists of an application of Parseval's identity for Fourier series. However, before we can attempt a solution to this problem, we must define the

notion of a simple closed curve, its length, and what we mean by the area of the region enclosed by it.

Curves, length and area

A parametrized curve γ is a mapping

$$\gamma : [a, b] \to \mathbb{R}^2.$$

The image of γ is a set of points in the plane which we call a **curve** and denote by Γ. The curve Γ is **simple** if it does not intersect itself, and **closed** if its two end-points coincide. In terms of the parametrization above, these two conditions translate into $\gamma(s_1) \neq \gamma(s_2)$ unless $s_1 = a$ and $s_2 = b$, in which case $\gamma(a) = \gamma(b)$. We may extend γ to a periodic function on $\mathbb{R}$ of period $b - a$, and think of γ as a function on the circle. We also always impose some smoothness on our curves by assuming that γ is of class C^1, and that its derivative γ' satisfies $\gamma'(s) \neq 0$. Altogether, these conditions guarantee that Γ has a well-defined tangent at each point, which varies continuously as the point on the curve varies. Moreover, the parametrization γ induces an orientation on Γ as the parameter s travels from a to b.

Any C^1 bijective mapping $s : [c, d] \to [a, b]$ gives rise to another parametrization of Γ by the formula

$$\eta(t) = \gamma(s(t)).$$

Clearly, the conditions that Γ be closed and simple are independent of the chosen parametrization. Also, we say that the two parametrizations γ and η are equivalent if $s'(t) > 0$ for all t; this means that η and γ induce the same orientation on the curve Γ. If, however, $s'(t) < 0$, then η reverses the orientation.

If Γ is parametrized by $\gamma(s) = (x(s), y(s))$, then the **length** of the curve Γ is defined by

$$\ell = \int_a^b |\gamma'(s)|\, ds = \int_a^b \left(x'(s)^2 + y'(s)^2\right)^{1/2}\, ds.$$

The length of Γ is a notion intrinsic to the curve, and does not depend on its parametrization. To see that this is indeed the case, suppose that $\gamma(s(t)) = \eta(t)$. Then, the change of variables formula and the chain rule imply that

$$\int_a^b |\gamma'(s)|\, ds = \int_c^d |\gamma'(s(t))|\, |s'(t)|\, dt = \int_c^d |\eta'(t)|\, dt,$$

as desired.

In the proof of the theorem below, we shall use a special type of parametrization for Γ. We say that γ is a **parametrization by arc-length** if $|\gamma'(s)| = 1$ for all s. This means that $\gamma(s)$ travels at a constant speed, and as a consequence, the length of Γ is precisely $b - a$. Therefore, after a possible additional translation, a parametrization by arc-length will be defined on $[0, \ell]$. Any curve admits a parametrization by arc-length (Exercise 1).

We now turn to the isoperimetric problem.

The attempt to give a precise formulation of the area $\mathcal{A}$ of the region enclosed by a simple closed curve Γ raises a number of tricky questions. In a variety of simple situations, it is evident that the area is given by the following familiar formula of the calculus:

$$\begin{aligned}\mathcal{A} &= \frac{1}{2}\left|\int_\Gamma (x\,dy - y\,dx)\right| \\ &= \frac{1}{2}\left|\int_a^b x(s)y'(s) - y(s)x'(s)\,ds\right|;\end{aligned} \tag{1}$$

see, for example, Exercise 3. Thus in formulating our result we shall adopt the easy expedient of taking (1) as our definition of area. This device allows us to give a quick and neat proof of the isoperimetric inequality. A listing of issues this simplification leaves unresolved can be found after the proof of the theorem.

Statement and proof of the isoperimetric inequality

Theorem 1.1 *Suppose that Γ is a simple closed curve in $\mathbb{R}^2$ of length ℓ, and let $\mathcal{A}$ denote the area of the region enclosed by this curve. Then*

$$\mathcal{A} \leq \frac{\ell^2}{4\pi},$$

with equality if and only if Γ is a circle.

The first observation is that we can rescale the problem. This means that we can change the units of measurement by a factor of $\delta > 0$ as follows. Consider the mapping of the plane $\mathbb{R}^2$ to itself, which sends the point (x, y) to $(\delta x, \delta y)$. A look at the formula defining the length of a curve shows that if Γ is of length ℓ, then its image under this mapping has length $\delta\ell$. So this operation magnifies or contracts lengths by a factor of δ depending on whether $\delta \geq 1$ or $\delta \leq 1$. Similarly, we see that

the mapping magnifies (or contracts) areas by a factor of δ^2. By taking $\delta = 2\pi/\ell$, we see that it suffices to prove that if $\ell = 2\pi$ then $\mathcal{A} \leq \pi$, with equality only if Γ is a circle.

Let $\gamma : [0, 2\pi] \to \mathbb{R}^2$ with $\gamma(s) = (x(s), y(s))$ be a parametrization by arc-length of the curve Γ, that is, $x'(s)^2 + y'(s)^2 = 1$ for all $s \in [0, 2\pi]$. This implies that

$$\frac{1}{2\pi}\int_0^{2\pi} (x'(s)^2 + y'(s)^2)\, ds = 1. \tag{2}$$

Since the curve is closed, the functions $x(s)$ and $y(s)$ are 2π-periodic, so we may consider their Fourier series

$$x(s) \sim \sum a_n e^{ins} \quad \text{and} \quad y(s) \sim \sum b_n e^{ins}.$$

Then, as we remarked in the later part of Section 2 of Chapter 2, we have

$$x'(s) \sim \sum a_n ine^{ins} \quad \text{and} \quad y'(s) \sim \sum b_n ine^{ins}.$$

Parseval's identity applied to (2) gives

$$\sum_{n=-\infty}^{\infty} |n|^2 \left(|a_n|^2 + |b_n|^2\right) = 1. \tag{3}$$

We now apply the bilinear form of Parseval's identity (Lemma 1.5, Chapter 3) to the integral defining $\mathcal{A}$. Since $x(s)$ and $y(s)$ are real-valued, we have $a_n = \overline{a_{-n}}$ and $b_n = \overline{b_{-n}}$, so we find that

$$\mathcal{A} = \frac{1}{2}\left|\int_0^{2\pi} x(s)y'(s) - y(s)x'(s)\, ds\right| = \pi\left|\sum_{n=-\infty}^{\infty} n\left(a_n\overline{b_n} - b_n\overline{a_n}\right)\right|.$$

We observe next that

$$|a_n\overline{b_n} - b_n\overline{a_n}| \leq 2\,|a_n|\,|b_n| \leq |a_n|^2 + |b_n|^2, \tag{4}$$

and since $|n| \leq |n|^2$, we may use (3) to get

$$\begin{aligned} \mathcal{A} &\leq \pi \sum_{n=-\infty}^{\infty} |n|^2 \left(|a_n|^2 + |b_n|^2\right) \\ &\leq \pi, \end{aligned}$$

as desired.

When $\mathcal{A} = \pi$, we see from the above argument that

$$x(s) = a_{-1}e^{-is} + a_0 + a_1 e^{is} \quad \text{and} \quad y(s) = b_{-1}e^{-is} + b_0 + b_1 e^{is}$$

because $|n| < |n|^2$ as soon as $|n| \geq 2$. We know that $x(s)$ and $y(s)$ are real-valued, so $a_{-1} = \overline{a_1}$ and $b_{-1} = \overline{b_1}$. The identity (3) implies that $2\left(|a_1|^2 + |b_1|^2\right) = 1$, and since we have equality in (4) we must have $|a_1| = |b_1| = 1/2$. We write

$$a_1 = \frac{1}{2}\, e^{i\alpha} \quad \text{and} \quad b_1 = \frac{1}{2}\, e^{i\beta}.$$

The fact that $1 = 2|a_1\overline{b_1} - \overline{a_1}b_1|$ implies that $|\sin(\alpha - \beta)| = 1$, hence $\alpha - \beta = k\pi/2$ where k is an odd integer. From this we find that

$$x(s) = a_0 + \cos(\alpha + s) \quad \text{and} \quad y(s) = b_0 \pm \sin(\alpha + s),$$

where the sign in $y(s)$ depends on the parity of $(k-1)/2$. In any case, we see that Γ is a circle, for which the case of equality obviously holds, and the proof of the theorem is complete.

The solution given above (due to Hurwitz in 1901) is indeed very elegant, but clearly leaves some important issues unanswered. We list these as follows. Suppose Γ is a simple closed curve.

(i) How is the "region enclosed by Γ" defined?

(ii) What is the geometric definition of the "area" of this region? Does this definition accord with (1)?

(iii) Can these results be extended to the most general class of simple closed curves relevant to the problem—those curves which are "rectifiable"—that is, those to which we can ascribe a finite length?

It turns out that the clarifications of the problems raised are connected to a number of other significant ideas in analysis. We shall return to these questions in succeeding books of this series.

2 Weyl's equidistribution theorem

We now apply ideas coming from Fourier series to a problem dealing with properties of irrational numbers. We begin with a brief discussion of congruences, a concept needed to understand our main theorem.

The reals modulo the integers

If x is a real number, we let $[x]$ denote the greatest integer less than or equal to x and call the quantity $[x]$ the **integer part** of x. The **fractional part** of x is then defined by $\langle x\rangle = x - [x]$. In particular, $\langle x\rangle \in [0,1)$ for every $x \in \mathbb{R}$. For example, the integer and fractional parts of 2.7 are 2 and 0.7, respectively, while the integer and fractional parts of -3.4 are -4 and 0.6, respectively.

We may define a relation on $\mathbb{R}$ by saying that the two numbers x and y are equivalent, or congruent, if $x - y \in \mathbb{Z}$. We then write

$$x = y \bmod \mathbb{Z} \quad \text{or} \quad x = y \bmod 1.$$

This means that we identify two real numbers if they differ by an integer. Observe that any real number x is congruent to a unique number in $[0,1)$ which is precisely $\langle x\rangle$, the fractional part of x. In effect, reducing a real number modulo $\mathbb{Z}$ means looking only at its fractional part and disregarding its integer part.

Now start with a real number $\gamma \neq 0$ and look at the sequence γ, 2γ, 3γ, $\ldots$. An intriguing question is to ask what happens to this sequence if we reduce it modulo $\mathbb{Z}$, that is, if we look at the sequence of fractional parts

$$\langle\gamma\rangle,\ \langle 2\gamma\rangle,\ \langle 3\gamma\rangle,\ \ldots .$$

Here are some simple observations:

(i) If γ is rational, then only finitely many numbers appearing in $\langle n\gamma\rangle$ are distinct.

(ii) If γ is irrational, then the numbers $\langle n\gamma\rangle$ are all distinct.

Indeed, for part (i), note that if $\gamma = p/q$, the first q terms in the sequence are

$$\langle p/q\rangle,\ \langle 2p/q\rangle,\ \ldots,\ \langle (q-1)p/q\rangle,\ \langle qp/q\rangle = 0.$$

The sequence then begins to repeat itself, since

$$\langle (q+1)p/q\rangle = \langle 1 + p/q\rangle = \langle p/q\rangle,$$

and so on. However, see Exercise 6 for a more refined result.

Also, for part (ii) assume that not all numbers are distinct. We therefore have $\langle n_1\gamma\rangle = \langle n_2\gamma\rangle$ for some $n_1 \neq n_2$; then $n_1\gamma - n_2\gamma \in \mathbb{Z}$, hence γ is rational, a contradiction.

In fact, it can be shown that if γ is irrational, then $\langle n\gamma\rangle$ is dense in the interval $[0,1)$, a result originally proved by Kronecker. In other words, the sequence $\langle n\gamma\rangle$ hits every sub-interval of $[0,1)$ (and hence it does so infinitely many times). We will obtain this fact as a corollary of a deeper theorem dealing with the uniform distribution of the sequence $\langle n\gamma\rangle$.

A sequence of numbers $\xi_1, \xi_2, \dots, \xi_n, \dots$ in $[0,1)$ is said to be **equidistributed** if for every interval $(a,b) \subset [0,1)$,

$$\lim_{N\to\infty} \frac{\#\{1 \leq n \leq N : \xi_n \in (a,b)\}}{N} = b - a$$

where $\#A$ denotes the cardinality of the finite set A. This means that for large N, the proportion of numbers ξ_n in (a,b) with $n \leq N$ is equal to the ratio of the length of the interval (a,b) to the length of the interval $[0,1)$. In other words, the sequence ξ_n sweeps out the whole interval evenly, and every sub-interval gets its fair share. Clearly, the ordering of the sequence is very important, as the next two examples illustrate.

EXAMPLE 1. The sequence

$$0,\ \frac{1}{2},\ 0,\ \frac{1}{3},\ \frac{2}{3},\ 0,\ \frac{1}{4},\ \frac{2}{4},\ \frac{3}{4},\ 0,\ \frac{1}{5},\ \frac{2}{5},\ \cdots$$

appears to be equidistributed since it passes over the interval $[0,1)$ very evenly. Of course this is not a proof, and the reader is invited to give one. For a somewhat related example, see Exercise 8 with $\sigma = 1/2$.

EXAMPLE 2. Let $\{r_n\}_{n=1}^{\infty}$ be *any* enumeration of the rationals in $[0,1)$. Then the sequence defined by

$$\xi_n = \begin{cases} r_{n/2} & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd,} \end{cases}$$

is not equidistributed since "half" of the sequence is at 0. Nevertheless, this sequence is obviously dense.

We now arrive at the main theorem of this section.

Theorem 2.1 *If γ is irrational, then the sequence of fractional parts $\langle\gamma\rangle, \langle 2\gamma\rangle, \langle 3\gamma\rangle, \dots$ is equidistributed in* $[0,1)$.

In particular, $\langle n\gamma\rangle$ is dense in $[0,1)$, and we get Kronecker's theorem as a corollary. In Figure 2 we illustrate the set of points $\langle\gamma\rangle, \langle 2\gamma\rangle, \langle 3\gamma\rangle, \dots, \langle N\gamma\rangle$ for three different values of N when $\gamma = \sqrt{2}$.

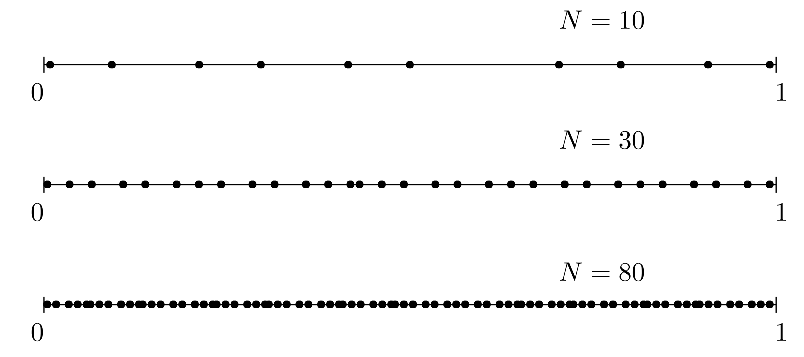


Figure 2. The sequence $\langle\gamma\rangle, \langle 2\gamma\rangle, \langle 3\gamma\rangle, \ldots, \langle N\gamma\rangle$ when $\gamma = \sqrt{2}$

Fix $(a, b) \subset [0, 1)$ and let $\chi_{(a,b)}(x)$ denote the characteristic function of the interval (a, b), that is, the function equal to 1 in (a, b) and 0 in $[0, 1) - (a, b)$. We may extend this function to $\mathbb{R}$ by periodicity (period 1), and still denote this extension by $\chi_{(a,b)}(x)$. Then, as a consequence of the definitions, we find that

$$\#\{1 \leq n \leq N : \langle n\gamma\rangle \in (a, b)\} = \sum_{n=1}^{N} \chi_{(a,b)}(n\gamma),$$

and the theorem can be reformulated as the statement that

$$\frac{1}{N}\sum_{n=1}^{N} \chi_{(a,b)}(n\gamma) \to \int_0^1 \chi_{(a,b)}(x)\, dx, \qquad \text{as } N \to \infty.$$

This step removes the difficulty of working with fractional parts and reduces the number theory to analysis.

The heart of the matter lies in the following result.

Lemma 2.2 *If f is continuous and periodic of period 1, and γ is irrational, then*

$$\frac{1}{N}\sum_{n=1}^{N} f(n\gamma) \to \int_0^1 f(x)\, dx \qquad \text{as } N \to \infty.$$

The proof of the lemma is divided into three steps.

Step 1. We first check the validity of the limit in the case when f is one of the exponentials $1, e^{2\pi ix}, \ldots, e^{2\pi ikx}, \ldots$. If $f = 1$, the limit

surely holds. If $f = e^{2\pi ikx}$ with $k \neq 0$, then the integral is 0. Since γ is irrational, we have $e^{2\pi ik\gamma} \neq 1$, therefore

$$\frac{1}{N}\sum_{n=1}^{N} f(n\gamma) = \frac{e^{2\pi ik\gamma}}{N}\frac{1-e^{2\pi ikN\gamma}}{1-e^{2\pi ik\gamma}},$$

which goes to 0 as $N \to \infty$.

Step 2. It is clear that if f and g satisfy the lemma, then so does $Af + Bg$ for any $A, B \in \mathbb{C}$. Therefore, the first step implies that the lemma is true for all trigonometric polynomials.

Step 3. Let $\epsilon > 0$. If f is any continuous periodic function of period 1, choose a trigonometric polynomial P so that $\sup_{x\in\mathbb{R}} |f(x) - P(x)| < \epsilon/3$ (this is possible by Corollary 5.4 in Chapter 2). Then, by step 1, for all large N we have

$$\left|\frac{1}{N}\sum_{n=1}^{N} P(n\gamma) - \int_0^1 P(x)\,dx\right| < \epsilon/3.$$

Therefore

$$\begin{aligned}\left|\frac{1}{N}\sum_{n=1}^{N} f(n\gamma) - \int_0^1 f(x)\,dx\right| &\leq \frac{1}{N}\sum_{n=1}^{N} |f(n\gamma) - P(n\gamma)| + \\ &\qquad + \left|\frac{1}{N}\sum_{n=1}^{N} P(n\gamma) - \int_0^1 P(x)\,dx\right| + \\ &\qquad + \int_0^1 |P(x) - f(x)|\,dx \\ &< \epsilon,\end{aligned}$$

and the lemma is proved.

Now we can finish the proof of the theorem. Choose two continuous periodic functions f_ϵ^+ and f_ϵ^- of period 1 which approximate $\chi_{(a,b)}(x)$ on $[0, 1)$ from above and below; both f_ϵ^+ and f_ϵ^- are bounded by 1 and agree with $\chi_{(a,b)}(x)$ except in intervals of total length 2ϵ (see Figure 3).
In particular, $f_\epsilon^-(x) \leq \chi_{(a,b)}(x) \leq f_\epsilon^+(x)$, and

$$b - a - 2\epsilon \leq \int_0^1 f_\epsilon^-(x)\,dx \quad \text{and} \quad \int_0^1 f_\epsilon^+(x)\,dx \leq b - a + 2\epsilon.$$

If $S_N = \frac{1}{N}\sum_{n=1}^{N} \chi_{(a,b)}(n\gamma)$, then we get

$$\frac{1}{N}\sum_{n=1}^{N} f_\epsilon^-(n\gamma) \leq S_N \leq \frac{1}{N}\sum_{n=1}^{N} f_\epsilon^+(n\gamma).$$

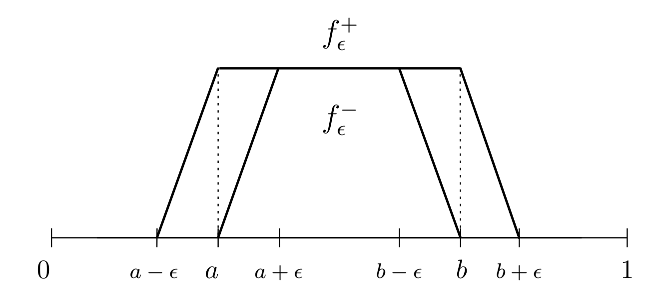


Figure 3. Approximations of $\chi_{(a,b)}(x)$

Therefore

$$b - a - 2\epsilon \le \liminf_{N\to\infty} S_N \quad \text{and} \quad \limsup_{N\to\infty} S_N \le b - a + 2\epsilon.$$

Since this is true for every $\epsilon > 0$, the limit $\lim_{N\to\infty} S_N$ exists and must equal $b - a$. This completes the proof of the equidistribution theorem.

This theorem has the following consequence.

Corollary 2.3 *The conclusion of Lemma 2.2 holds for every function f which is Riemann integrable in $[0,1]$, and periodic of period 1.*

Proof. Assume f is real-valued, and consider a partition of the interval $[0,1]$, say $0 = x_0 < x_1 < \cdots < x_N = 1$. Next, define $f_U(x) = \sup_{x_{j-1}\le y\le x_j} f(y)$ if $x \in [x_{j-1}, x_j)$ and $f_L(x) = \inf_{x_{j-1}\le y\le x_j} f(y)$ for $x \in (x_{j-1}, x_j)$. Then clearly $f_L \le f \le f_U$ and

$$\int_0^1 f_L(x)\,dx \le \int_0^1 f(x)\,dx \le \int_0^1 f_U(x)\,dx.$$

Moreover, by making the partition sufficiently fine we can guarantee that for a given $\epsilon > 0$,

$$\int_0^1 f_U(x)\,dx - \int_0^1 f_L(x)\,dx \le \epsilon.$$

However,

$$\frac{1}{N}\sum_{n=1}^{N} f_L(n\gamma) \to \int_0^1 f_L(x)\,dx$$

by the theorem, because each f_L is a finite linear combination of characteristic functions of intervals; similarly we have

$$\frac{1}{N}\sum_{n=1}^{N} f_U(n\gamma) \to \int_0^1 f_U(x)\,dx.$$

From these two assertions we can conclude the proof of the corollary by using the previous approximation argument.

There is an interesting interpretation of the lemma and its corollary, in terms of a simple dynamical system. In this example, the underlying space is the circle parametrized by the angle θ. We also consider a mapping of this space to itself: here, we choose a rotation ρ of the circle by the angle $2\pi\gamma$, that is, the transformation $\rho : \theta \mapsto \theta + 2\pi\gamma$.

We want next to consider how this space, with its underlying action ρ, evolves in time. In other words, we wish to consider the iterates of ρ, namely ρ, ρ^2, ρ^3, $\ldots$, ρ^n where

$$\rho^n = \rho \circ \rho \circ \cdots \circ \rho : \theta \mapsto \theta + 2\pi n\gamma,$$

and where we think of the action ρ^n taking place at the time $t = n$.

To each Riemann integrable function f on the circle, we can also associate the corresponding effects of the rotation ρ, and obtain a sequence of functions

$$f(\theta),\ f(\rho(\theta)),\ f(\rho^2(\theta)),\ \ldots,\ f(\rho^n(\theta)),\ \ldots$$

with $f(\rho^n(\theta)) = f(\theta + 2\pi n\gamma)$. In this special context, the **ergodicity** of this system is then the statement that the "time average"

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\rho^n(\theta))$$

exists for each θ and equals the "space average"

$$\frac{1}{2\pi}\int_0^{2\pi} f(\theta)\,d\theta,$$

whenever γ is irrational. In fact, this assertion is merely a rephrasing of Corollary 2.3, once we make the change of variables $\theta = 2\pi x$.

Returning to the problem of equidistributed sequences, we observe that the proof of Theorem 2.1 gives the following characterization.

Weyl's criterion. *A sequence of real numbers* $\xi_1, \xi_2 \ldots$ *in* $[0,1)$ *is equidistributed if and only if for all integers* $k \neq 0$ *one has*

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i k \xi_n} \to 0, \qquad \text{as } N \to \infty.$$

One direction of this theorem was in effect proved above, and the converse can be found in Exercise 7. In particular, we find that to understand the equidistributive properties of a sequence ξ_n, it suffices to estimate the size of the corresponding "exponential sum" $\sum_{n=1}^{N} e^{2\pi i k \xi_n}$. For example, it can be shown using Weyl's criterion that the sequence $\langle n^2\gamma\rangle$ is equidistributed whenever γ is irrational. This and other examples can be found in Exercises 8, and 9; also Problems 2, and 3.

As a last remark, we mention a nice geometric interpretation of the distribution properties of $\langle n\gamma\rangle$. Suppose that the sides of a square are reflecting mirrors and that a ray of light leaves a point inside the square. What kind of path will the light trace out?

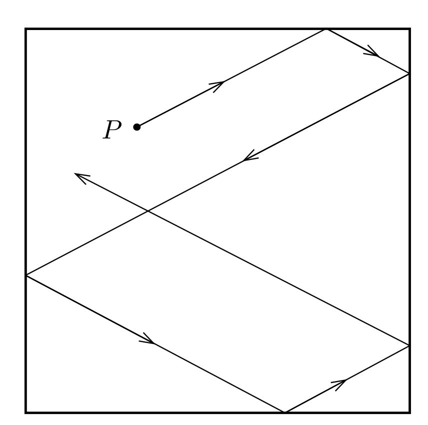


Figure 4. Reflection of a ray of light in a square

To solve this problem, the main idea is to consider the grid of the plane formed by successively reflecting the initial square across its sides. With an appropriate choice of axis, the path traced by the light in the square corresponds to the straight line $P + (t, \gamma t)$ in the plane. As a result, the reader may observe that the path will be either closed and periodic, or it will be dense in the square. The first of these situations

will happen if and only if the slope γ of the initial direction of the light (determined with respect to one of the sides of the square) is rational. In the second situation, when γ is irrational, the density follows from Kronecker's theorem. What stronger conclusion does one get from the equidistribution theorem?

3 A continuous but nowhere differentiable function

There are many obvious examples of continuous functions that are not differentiable at one point, say $f(x) = |x|$. It is almost as easy to construct a continuous function that is not differentiable at any given finite set of points, or even at appropriate sets containing countably many points. A more subtle problem is whether there exists a continuous function that is *nowhere* differentiable. In 1861, Riemann guessed that the function defined by

$$R(x) = \sum_{n=1}^{\infty} \frac{\sin(n^2 x)}{n^2} \tag{5}$$

was nowhere differentiable. He was led to consider this function because of its close connection to the theta function which will be introduced in Chapter 5. Riemann never gave a proof, but mentioned this example in one of his lectures. This triggered the interest of Weierstrass who, in an attempt to find a proof, came across the first example of a continuous but nowhere differentiable function. Say $0 < b < 1$ and a is an integer > 1. In 1872 he proved that if $ab > 1 + 3\pi/2$, then the function

$$W(x) = \sum_{n=1}^{\infty} b^n \cos(a^n x)$$

is nowhere differentiable.

But the story is not complete without a final word about Riemann's original function. In 1916 Hardy showed that R is not differentiable at all irrational multiples of π, and also at certain rational multiples of π. However, it was not until much later, in 1969, that Gerver completely settled the problem, first by proving that the function R *is* actually differentiable at all the rational multiples of π of the form $\pi p/q$ with p and q odd integers, and then by showing that R is not differentiable in all of the remaining cases.

In this section, we prove the following theorem.

Theorem 3.1 *If* $0 < \alpha < 1$*, then the function*

$$f_\alpha(x) = f(x) = \sum_{n=0}^{\infty} 2^{-n\alpha} e^{i2^n x}$$

is continuous but nowhere differentiable.

The continuity is clear because of the absolute convergence of the series. The crucial property of f which we need is that it has many vanishing Fourier coefficients. A Fourier series that skips many terms, like the one given above, or like $W(x)$, is called a **lacunary Fourier series**.

The proof of the theorem is really the story of three methods of summing a Fourier series. First, there is the ordinary convergence in terms of the partial sums $S_N(g) = g * D_N$. Next, there is Cesàro summability $\sigma_N(g) = g * F_N$, with F_N the Fejér kernel. A third method, clearly connected with the second, involves the **delayed means** defined by

$$\triangle_N(g) = 2\sigma_{2N}(g) - \sigma_N(g).$$

Hence $\triangle_N(g) = g * [2F_{2N} - F_N]$. These methods can best be visualized as in Figure 5.

Suppose $g(x) \sim \sum a_n e^{inx}$. Then:

- S_N arises by multiplying the term $a_n e^{inx}$ by 1 if $|n| \leq N$, and 0 if $|n| > N$.
- σ_N arises by multiplying $a_n e^{inx}$ by $1 - |n|/N$ for $|n| \leq N$ and 0 for $|n| > N$.
- $\triangle_N$ arises by multiplying $a_n e^{inx}$ by 1 if $|n| \leq N$, by $2(1 - |n|/(2N))$ for $N \leq |n| \leq 2N$, and 0 for $|n| > 2N$.

For example, note that

$$\begin{aligned}
\sigma_N(g)(x) &= \frac{S_0(g)(x) + S_1(g)(x) + \cdots + S_{N-1}(g)(x)}{N} \\
&= \frac{1}{N} \sum_{\ell=0}^{N-1} \sum_{|k| \leq \ell} a_k e^{ikx} \\
&= \frac{1}{N} \sum_{|n| \leq N} (N - |n|) a_n e^{inx} \\
&= \sum_{|n| \leq N} \left(1 - \frac{|n|}{N}\right) a_n e^{inx}.
\end{aligned}$$

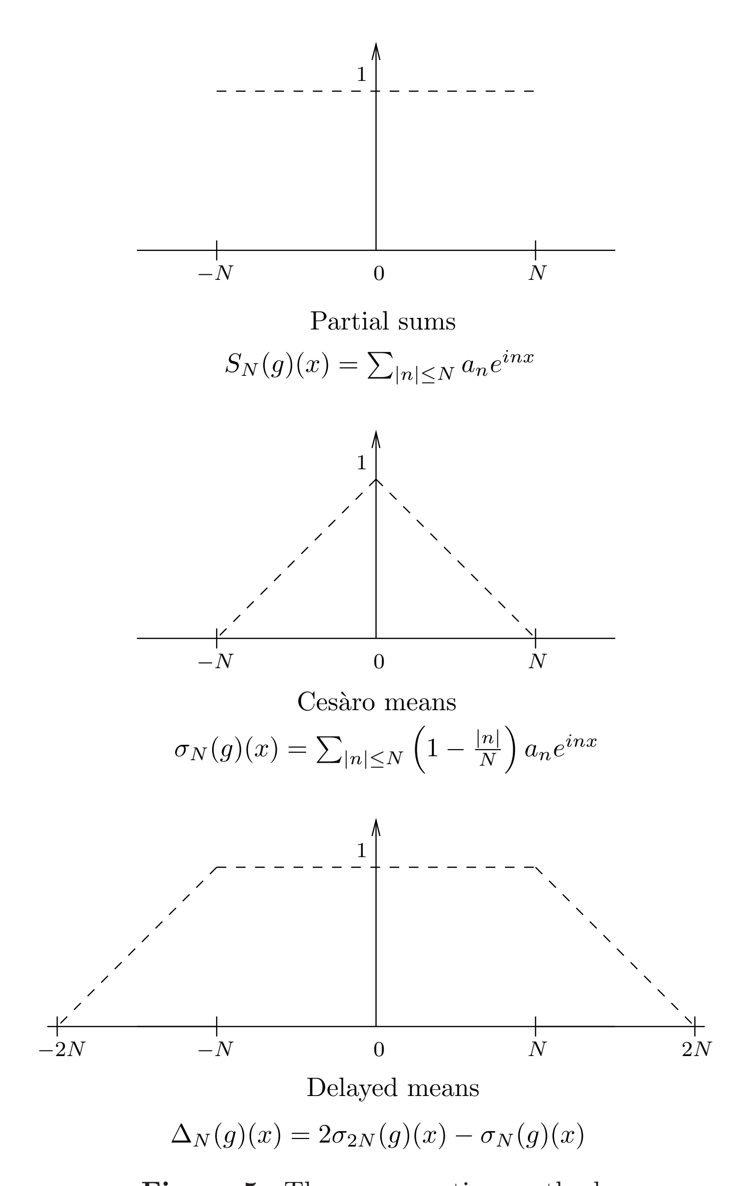


Figure 5. Three summation methods

The proof of the other assertion is similar.

The delayed means have two important features. On the one hand, their properties are closely related to the (good) features of the Cesàro means. On the other hand, for series that have lacunary properties like those of f, the delayed means are essentially equal to the partial sums. In particular, note that for our function $f = f_\alpha$

$$S_N(f) = \triangle_{N'}(f), \tag{6}$$

where N' is the largest integer of the form 2^k with $N' \leq N$. This is clear by examining Figure 5 and the definition of f.

We turn to the proof of the theorem proper and argue by contradiction; that is, we assume that $f'(x_0)$ exists for some x_0.

Lemma 3.2 *Let g be any continuous function that is differentiable at x_0. Then, the Cesàro means satisfy $\sigma_N(g)'(x_0) = O(\log N)$, therefore*

$$\triangle_N(g)'(x_0) = O(\log N).$$

Proof. First we have

$$\sigma_N(g)'(x_0) = \int_{-\pi}^{\pi} F_N'(x_0 - t)g(t)\,dt = \int_{-\pi}^{\pi} F_N'(t)g(x_0 - t)\,dt,$$

where F_N is the Fejér kernel. Since F_N is periodic, we have $\int_{-\pi}^{\pi} F_N'(t)dt=0$ and this implies that

$$\sigma_N(g)'(x_0) = \int_{-\pi}^{\pi} F_N'(t)[g(x_0 - t) - g(x_0)]\,dt.$$

From the assumption that g is differentiable at x_0 we get

$$|\sigma_N(g)'(x_0)| \leq C\int_{-\pi}^{\pi} |F_N'(t)|\,|t|\,dt.$$

Now observe that F_N' satisfies the two estimates

$$|F_N'(t)| \leq AN^2 \quad \text{and} \quad |F_N'(t)| \leq \frac{A}{|t|^2}.$$

For the first inequality, recall that F_N is a trigonometric polynomial of degree N whose coefficients are bounded by 1. Therefore, F_N' is a trigonometric polynomial of degree N whose coefficients are no bigger than N. Hence $|F'(t)| \leq (2N + 1)N \leq AN^2$.

For the second inequality, we recall that

$$F_N(t) = \frac{1}{N}\frac{\sin^2(Nt/2)}{\sin^2(t/2)}.$$

Differentiating this expression, we get two terms:

$$\frac{\sin(Nt/2)\cos(Nt/2)}{\sin^2(t/2)} - \frac{1}{N}\frac{\cos(t/2)\sin^2(Nt/2)}{\sin^3(t/2)}.$$

If we then use the facts that $|\sin(Nt/2)| \leq CN|t|$ and $|\sin(t/2)| \geq c|t|$ (for $|t| \leq \pi$), we get the desired estimates for $F_N'(t)$.

Using all of these estimates we find that

$$\begin{aligned}
|\sigma_N(g)'(x_0)| &\leq C\int_{|t|\geq 1/N} |F_N'(t)|\,|t|\,dt + C\int_{|t|\leq 1/N} |F_N'(t)|\,|t|\,dt \\
&\leq CA\int_{|t|\geq 1/N} \frac{dt}{|t|} + CAN\int_{|t|\leq 1/N} dt \\
&= O(\log N) + O(1) \\
&= O(\log N).
\end{aligned}$$

The proof of the lemma is complete once we invoke the definition of $\triangle_N(g)$.

Lemma 3.3 *If* $2N = 2^n$, *then*

$$\triangle_{2N}(f) - \triangle_N(f) = 2^{-n\alpha}e^{i2^n x}.$$

This follows from our previous observation (6) because $\triangle_{2N}(f) = S_{2N}(f)$ and $\triangle_N(f) = S_N(f)$.

Now, by the first lemma we have

$$\triangle_{2N}(f)'(x_0) - \triangle_N(f)'(x_0) = O(\log N),$$

and the second lemma also implies

$$|\triangle_{2N}(f)'(x_0) - \triangle_N(f)'(x_0)| = 2^{n(1-\alpha)} \geq cN^{1-\alpha}.$$

This is the desired contradiction since $N^{1-\alpha}$ grows faster than $\log N$.

A few additional remarks about our function $f_\alpha(x) = \sum_{n=0}^{\infty} 2^{-n\alpha}e^{i2^n x}$ are in order.

This function is complex-valued as opposed to the examples R and W above, and so the nowhere differentiability of f_α does not imply the same property for its real and imaginary parts. However, a small modification of our proof shows that, in fact, the real part of f_α,

$$\sum_{n=0}^{\infty} 2^{-n\alpha} \cos 2^n x,$$

as well as its imaginary part, are both nowhere differentiable. To see this, observe first that by the same proof, Lemma 3.2 has the following generalization: if g is a continuous function which is differentiable at x_0, then

$$\triangle_N(g)'(x_0+h) = O(\log N) \qquad \text{whenever } |h| \leq c/N.$$

We then proceed with $F(x) = \sum_{n=0}^{\infty} 2^{-n\alpha} \cos 2^n x$, noting as above that $\triangle_{2N}(F) - \triangle_N(F) = 2^{-n\alpha}\cos 2^n x$; as a result, assuming that F is differentiable at x_0, we get that

$$|2^{n(1-\alpha)} \sin(2^n(x_0+h))| = O(\log N)$$

when $2N = 2^n$, and $|h| \leq c/N$. To get a contradiction, we need only choose h so that $|\sin(2^n(x_0+h))| = 1$; this is accomplished by setting δ equal to the distance from $2^n x_0$ to the nearest number of the form $(k+1/2)\pi$, $k \in \mathbb{Z}$ (so $\delta \leq \pi/2$), and taking $h = \pm\delta/2^n$.

Clearly, when $\alpha > 1$ the function f_α is continuously differentiable since the series can be differentiated term by term. Finally, the nowhere differentiability we have proved for $\alpha < 1$ actually extends to $\alpha = 1$ by a suitable refinement of the argument (see Problem 8 in Chapter 5). In fact, using these more elaborate methods one can also show that the Weierstrass function W is nowhere differentiable if $ab \geq 1$.

4 The heat equation on the circle

As a final illustration, we return to the original problem of heat diffusion considered by Fourier.

Suppose we are given an initial temperature distribution at $t = 0$ on a ring and that we are asked to describe the temperature at points on the ring at times $t > 0$.

The ring is modeled by the unit circle. A point on this circle is described by its angle $\theta = 2\pi x$, where the variable x lies between 0 and 1. If $u(x,t)$ denotes the temperature at time t of a point described by the

angle θ, then considerations similar to the ones given in Chapter 1 show that u satisfies the differential equation

$$\frac{\partial u}{\partial t} = c\,\frac{\partial^2 u}{\partial x^2}. \tag{7}$$

The constant c is a positive physical constant which depends on the material of which the ring is made (see Section 2.1 in Chapter 1). After rescaling the time variable, we may assume that $c = 1$. If f is our initial data, we impose the condition

$$u(x, 0) = f(x).$$

To solve the problem, we separate variables and look for special solutions of the form

$$u(x, t) = A(x)B(t).$$

Then inserting this expression for u into the heat equation we get

$$\frac{B'(t)}{B(t)} = \frac{A''(x)}{A(x)}.$$

Both sides are therefore constant, say equal to λ. Since A must be periodic of period 1, we see that the only possibility is $\lambda = -4\pi^2 n^2$, where $n \in \mathbb{Z}$. Then A is a linear combination of the exponentials $e^{2\pi inx}$ and $e^{-2\pi inx}$, and $B(t)$ is a multiple of $e^{-4\pi^2 n^2 t}$. By superposing these solutions, we are led to

$$u(x, t) = \sum_{n=-\infty}^{\infty} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx}, \tag{8}$$

where, setting $t = 0$, we see that $\{a_n\}$ are the Fourier coefficients of f.

Note that when f is Riemann integrable, the coefficients a_n are bounded, and since the factor $e^{-4\pi^2 n^2 t}$ tends to zero extremely fast, the series defining u converges. In fact, in this case, u is twice differentiable and solves equation (7).

The natural question with regard to the boundary condition is the following: do we have $u(x, t) \to f(x)$ as t tends to 0, and in what sense? A simple application of the Parseval identity shows that this limit holds in the mean square sense (Exercise 11). For a better understanding of the properties of our solution (8), we write it as

$$u(x, t) = (f * H_t)(x),$$

where H_t is the **heat kernel for the circle**, given by

$$H_t(x) = \sum_{n=-\infty}^{\infty} e^{-4\pi^2 n^2 t} e^{2\pi i n x}, \tag{9}$$

and where the convolution for functions with period 1 is defined by

$$(f * g)(x) = \int_0^1 f(x-y) g(y)\, dy.$$

An analogy between the heat kernel and the Poisson kernel (of Chapter 2) is given in Exercise 12. However, unlike in the case of the Poisson kernel, there is no elementary formula for the heat kernel. Nevertheless, it turns out that it is a good kernel (in the sense of Chapter 2). The proof is not obvious and requires the use of the celebrated Poisson summation formula, which will be taken up in Chapter 5. As a corollary, we will also find that H_t is everywhere positive, a fact that is also not obvious from its defining expression (9). We can, however, give the following heuristic argument for the positivity of H_t. Suppose that we begin with an initial temperature distribution f which is everywhere ≤ 0. Then it is physically reasonable to expect $u(x,t) \leq 0$ for all t since heat travels from hot to cold. Now

$$u(x,t) = \int_0^1 f(x-y) H_t(y)\, dy.$$

If H_t is negative for some x_0, then we may choose $f \leq 0$ supported near x_0, and this would imply $u(x_0, t) > 0$, which is a contradiction.

5 Exercises

1. Let $\gamma : [a,b] \to \mathbb{R}^2$ be a parametrization for the closed curve Γ.

(a) Prove that γ is a parametrization by arc-length if and only if the length of the curve from $\gamma(a)$ to $\gamma(s)$ is precisely $s - a$, that is,

$$\int_a^s |\gamma'(t)|\, dt = s - a.$$

(b) Prove that any curve Γ admits a parametrization by arc-length. [Hint: If η is any parametrization, let $h(s) = \int_a^s |\eta'(t)|\, dt$ and consider $\gamma = \eta \circ h^{-1}$.]

2. Suppose $\gamma : [a,b] \to \mathbb{R}^2$ is a parametrization for a closed curve Γ, with $\gamma(t) = (x(t), y(t))$.

(a) Show that

$$\frac{1}{2}\int_a^b (x(s)y'(s) - y(s)x'(s))\,ds = \int_a^b x(s)y'(s)\,ds = -\int_a^b y(s)x'(s)\,ds.$$

(b) Define the **reverse parametrization** of γ by $\gamma^- : [a,b] \to \mathbb{R}^2$ with $\gamma^-(t) = \gamma(b+a-t)$. The image of γ^- is precisely Γ, except that the points $\gamma^-(t)$ and $\gamma(t)$ travel in opposite directions. Thus γ^- "reverses" the orientation of the curve. Prove that

$$\int_\gamma (x\,dy - y\,dx) = -\int_{\gamma^-} (x\,dy - y\,dx).$$

In particular, we may assume (after a possible change in orientation) that

$$\mathcal{A} = \frac{1}{2}\int_a^b (x(s)y'(s) - y(s)x'(s))\,ds = \int_a^b x(s)y'(s)\,ds.$$

3. Suppose Γ is a curve in the plane, and that there exists a set of coordinates x and y so that the x-axis divides the curve into the union of the graph of two continuous functions $y = f(x)$ and $y = g(x)$ for $0 \le x \le 1$, and with $f(x) \ge g(x)$ (see Figure 6). Let Ω denote the region between the graphs of these two functions:

$$\Omega = \{(x,y) : \; 0 \le x \le 1 \;\text{ and }\; g(x) \le y \le f(x)\}.$$

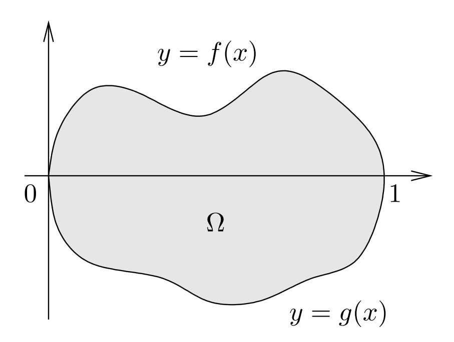


Figure 6. Simple version of the area formula

With the familiar interpretation that the integral $\int h(x)\,dx$ gives the area under the graph of the function h, we see that the area of Ω is $\int_0^1 f(x)\,dx -$

$\int_0^1 g(x)\,dx$. Show that this definition coincides with the area formula $\mathcal{A}$ given in the text, that is,

$$\int_0^1 f(x)\,dx - \int_0^1 g(x)\,dx = \left| - \int_\Gamma y\,dx \right| = \mathcal{A}.$$

Also, note that if the orientation of the curve is chosen so that Ω "lies to the left" of Γ, then the above formula holds without the absolute value signs.

This formula generalizes to any set that can be written as a finite union of domains like Ω above.

4. Observe that with the definition of ℓ and $\mathcal{A}$ given in the text, the isoperimetric inequality continues to hold (with the same proof) even when Γ is not simple.

Show that this stronger version of the isoperimetric inequality is equivalent to Wirtinger's inequality, which says that if f is 2π-periodic, of class C^1, and satisfies $\int_0^{2\pi} f(t)\,dt = 0$, then

$$\int_0^{2\pi} |f(t)|^2\,dt \leq \int_0^{2\pi} |f'(t)|^2\,dt$$

with equality if and only if $f(t) = A\sin t + B\cos t$ (Exercise 11, Chapter 3).
[Hint: In one direction, note that if the length of the curve is 2π and γ is an appropriate arc-length parametrization, then

$$2(\pi - \mathcal{A}) = \int_0^{2\pi} \left[x'(s) + y(s)\right]^2\,ds + \int_0^{2\pi} (y'(s)^2 - y(s)^2)\,ds.$$

A change of coordinates will guarantee $\int_0^{2\pi} y(s)\,ds = 0$. For the other direction, start with a real-valued f satisfying all the hypotheses of Wirtinger's inequality, and construct g, 2π-periodic and so that the term in brackets above vanishes.]

5. Prove that the sequence $\{\gamma_n\}_{n=1}^\infty$, where γ_n is the fractional part of

$$\left(\frac{1+\sqrt{5}}{2}\right)^n,$$

is not equidistributed in $[0,1]$.
[Hint: Show that $U_n = \left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n$ is the solution of the difference equation $U_{r+1} = U_r + U_{r-1}$ with $U_0 = 2$ and $U_1 = 1$. The U_n satisfy the same difference equation as the Fibonacci numbers.]

6. Let $\theta = p/q$ be a rational number where p and q are relatively prime integers (that is, θ is in lowest form). We assume without loss of generality that $q > 0$. Define a sequence of numbers in $[0,1)$ by $\xi_n = \langle n\theta \rangle$ where $\langle \cdot \rangle$ denotes the

fractional part. Show that the sequence $\{\xi_1, \xi_2, \ldots\}$ is equidistributed on the points of the form

$$0,\ 1/q,\ 2/q,\ \ldots,\ (q-1)/q.$$

In fact, prove that for any $0 \le a < q$, one has

$$\frac{\#\{n:\ 1 \le n \le N,\ \langle n\theta\rangle = a/q\}}{N} = \frac{1}{q} + O\left(\frac{1}{N}\right).$$

[Hint: For each integer $k \ge 0$, there exists a unique integer n with $kq \le n < (k+1)q$ and so that $\langle n\theta\rangle = a/q$. Why can one assume $k = 0$? Prove the existence of n by using the fact[1] that if p and q are relatively prime, there exist integers x, y such that $xp + yq = 1$. Next, divide N by q with remainder, that is, write $N = \ell q + r$ where $0 \le \ell$ and $0 \le r < q$. Establish the inequalities

$$\ell \le \#\{n:\ 1 \le n \le N,\ \langle n\theta\rangle = a/q\} \le \ell + 1.]$$

7. Prove the second part of Weyl's criterion: if a sequence of numbers $\xi_1, \xi_2, \ldots$ in $[0,1)$ is equidistributed, then for all $k \in \mathbb{Z} - \{0\}$

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i k\xi_n} \to 0 \qquad \text{as } N \to \infty.$$

[Hint: It suffices to show that $\frac{1}{N}\sum_{n=1}^{N} f(\xi_n) \to \int_0^1 f(x)\,dx$ for all continuous f. Prove this first when f is the characteristic function of an interval.]

8. Show that for any $a \ne 0$, and σ with $0 < \sigma < 1$, the sequence $\langle an^\sigma\rangle$ is equidistributed in $[0,1)$.

[Hint: Prove that $\sum_{n=1}^{N} e^{2\pi i b n^\sigma} = O(N^\sigma) + O(N^{1-\sigma})$ if $b \ne 0$.] In fact, note the following

$$\sum_{n=1}^{N} e^{2\pi i b n^\sigma} - \int_1^N e^{2\pi i b x^\sigma}\,dx = O\left(\sum_{n=1}^{N} n^{-1+\sigma}\right).$$

9. In contrast with the result in Exercise 8, prove that $\langle a \log n\rangle$ is *not* equidistributed for any a.

[Hint: Compare the sum $\sum_{n=1}^{N} e^{2\pi i b \log n}$ with the corresponding integral.]

10. Suppose that f is a periodic function on $\mathbb{R}$ of period 1, and $\{\xi_n\}$ is a sequence which is equidistributed in $[0,1)$. Prove that:

[1] The elementary results in arithmetic used in this exercise can be found at the beginning of Chapter 8.

(a) If f is continuous and satisfies $\int_0^1 f(x)\,dx = 0$, then

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x+\xi_n) = 0 \qquad \text{uniformly in } x.$$

[Hint: Establish this result first for trigonometric polynomials.]

(b) If f is merely integrable on $[0,1]$ and satisfies $\int_0^1 f(x)\,dx = 0$, then

$$\lim_{N\to\infty} \int_0^1 \left| \frac{1}{N}\sum_{n=1}^{N} f(x+\xi_n) \right|^2 dx = 0.$$

11. Show that if $u(x,t) = (f * H_t)(x)$ where H_t is the heat kernel, and f is Riemann integrable, then

$$\int_0^1 |u(x,t) - f(x)|^2\,dx \to 0 \qquad \text{as } t \to 0.$$

12. A change of variables in (8) leads to the solution

$$u(\theta,\tau) = \sum a_n e^{-n^2\tau} e^{in\theta} = (f * h_\tau)(\theta)$$

of the equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial \theta^2} \qquad \text{with } 0 \le \theta \le 2\pi \text{ and } \tau > 0,$$

with boundary condition $u(\theta,0) = f(\theta) \sim \sum a_n e^{in\theta}$. Here $h_\tau(\theta) = \sum_{n=-\infty}^{\infty} e^{-n^2\tau} e^{in\theta}$. This version of the heat kernel on $[0,2\pi]$ is the analogue of the Poisson kernel, which can be written as $P_r(\theta) = \sum_{n=-\infty}^{\infty} e^{-|n|\tau} e^{in\theta}$ with $r = e^{-\tau}$ (and so $0 < r < 1$ corresponds to $\tau > 0$).

13. The fact that the kernel $H_t(x)$ is a good kernel, hence $u(x,t) \to f(x)$ at each point of continuity of f, is not easy to prove. This will be shown in the next chapter. However, one can prove directly that $H_t(x)$ is "peaked" at $x = 0$ as $t \to 0$ in the following sense:

(a) Show that $\int_{-1/2}^{1/2} |H_t(x)|^2\,dx$ is of the order of magnitude of $t^{-1/2}$ as $t \to 0$. More precisely, prove that $t^{1/2}\int_{-1/2}^{1/2} |H_t(x)|^2\,dx$ converges to a non-zero limit as $t \to 0$.

(b) Prove that $\int_{-1/2}^{1/2} x^2 |H_t(x)|^2\,dx = O(t^{1/2})$ as $t \to 0$.

[Hint: For (a) compare the sum $\sum_{-\infty}^{\infty} e^{-cn^2 t}$ with the integral $\int_{-\infty}^{\infty} e^{-cx^2 t}\,dx$ where $c > 0$. For (b) use $x^2 \le C(\sin \pi x)^2$ for $-1/2 \le x \le 1/2$, and apply the mean value theorem to $e^{-cx^2 t}$.]

6 Problems

1.* This problem explores another relationship between the geometry of a curve and Fourier series. The diameter of a closed curve Γ parametrized by $\gamma(t) = (x(t), y(t))$ on $[-\pi, \pi]$ is defined by

$$d = \sup_{P,\, Q \in \Gamma} |P - Q| = \sup_{t_1,\, t_2 \in [-\pi,\pi]} |\gamma(t_1) - \gamma(t_2)|.$$

If a_n is the nth Fourier coefficient of $\gamma(t) = x(t) + iy(t)$ and ℓ denotes the length of Γ, then

(a) $2|a_n| \le d$ for all $n \neq 0$.

(b) $\ell \le \pi d$, whenever Γ is convex.

Property (a) follows from the fact that $2a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi}[\gamma(t) - \gamma(t + \pi/n)]e^{-int}\,dt$.

The equality $\ell = \pi d$ is satisfied when Γ is a circle, but surprisingly, this is not the only case. In fact, one finds that $\ell = \pi d$ is equivalent to $2|a_1| = d$. We re-parametrize γ so that for each t in $[-\pi, \pi]$ the tangent to the curve makes an angle t with the y-axis. Then, if $a_1 = 1$ we have

$$\gamma'(t) = ie^{it}(1 + r(t)),$$

where r is a real-valued function which satisfies $r(t) + r(t + \pi) = 0$, and $|r(t)| \le 1$. Figure 7 (a) shows the curve obtained by setting $r(t) = \cos 5t$. Also, Figure 7 (b) consists of the curve where $r(t) = h(3t)$, with $h(s) = -1$ if $-\pi \le s \le 0$ and $h(s) = 1$ if $0 < s < \pi$. This curve (which is only piecewise of class C^1) is known as the Reuleaux triangle and is the classical example of a convex curve of constant width which is not a circle.

2.* Here we present an estimate of Weyl which leads to some interesting results.

(a) Let $S_N = \sum_{n=1}^{N} e^{2\pi i f(n)}$. Show that for $H \le N$, one has

$$|S_N|^2 \le c\,\frac{N}{H}\sum_{h=0}^{H}\left|\sum_{n=1}^{N-h} e^{2\pi i (f(n+h) - f(n))}\right|,$$

for some constant $c > 0$ independent of N, H, and f.

(b) Use this estimate to show that the sequence $\langle n^2\gamma\rangle$ is equidistributed in $[0, 1)$ whenever γ is irrational.

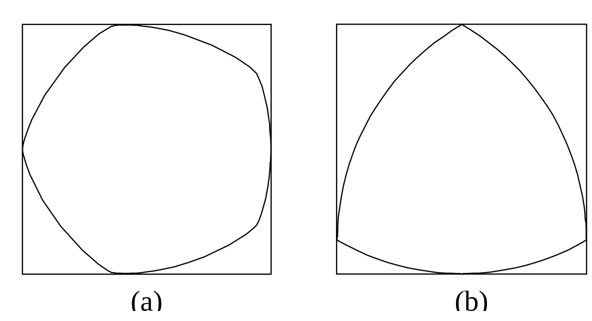


Figure 7. Some curves with maximal length for a given diameter

(c) More generally, show that if $\{\xi_n\}$ is a sequence of real numbers so that for all positive integers h the difference $\langle \xi_{n+h} - \xi_n \rangle$ is equidistributed in $[0, 1)$, then $\langle \xi_n \rangle$ is also equidistributed in $[0, 1)$.

(d) Suppose that $P(x) = c_n x^n + \cdots + c_0$ is a polynomial with real coefficients, where at least one of $c_1, \ldots, c_n$ is irrational. Then the sequence $\langle P(n) \rangle$ is equidistributed in $[0, 1)$.

[Hint: For (a), let $a_n = e^{2\pi i f(n)}$ when $1 \leq n \leq N$ and 0 otherwise. Then write $H \sum_n a_n = \sum_{k=1}^{H} \sum_n a_{n+k}$ and apply the Cauchy-Schwarz inequality. For (b), note that $(n + h)^2\gamma - n^2\gamma = 2nh\gamma + h^2\gamma$, and use the fact that for each integer h, the sequence $\langle 2nh\gamma \rangle$ is equidistributed. Finally, to prove (d), assume first that $P(x) = Q(x) + c_1 x + c_0$ where c_1 is irrational, and estimate the exponential sum $\sum_{n=1}^{N} e^{2\pi i k P(n)}$. Then, argue by induction on the highest degree term which has an irrational coefficient, and use part (c).]

3.* If $\sigma > 0$ is not an integer and $a \neq 0$, then $\langle a n^\sigma \rangle$ is equidistributed in $[0, 1)$. See also Exercise 8.

4. An elementary construction of a continuous but nowhere differentiable function is obtained by "piling up singularities," as follows.

On $[-1, 1]$ consider the function

$$\varphi(x) = |x|$$

and extend φ to $\mathbb{R}$ by requiring it to be periodic of period 2. Clearly, φ is continuous on $\mathbb{R}$ and $|\varphi(x)| \leq 1$ for all x so the function f defined by

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x)$$

is continuous on $\mathbb{R}$.

(a) Fix $x_0 \in \mathbb{R}$. For every positive integer m, let $\delta_m = \pm\frac{1}{2}4^{-m}$ where the sign is chosen so that no integer lies in between $4^m x_0$ and $4^m(x_0 + \delta_m)$. Consider the quotient

$$\gamma_n = \frac{\varphi(4^n(x_0 + \delta_m)) - \varphi(4^n x_0)}{\delta_m}.$$

Prove that if $n > m$, then $\gamma_n = 0$, and for $0 \leq n \leq m$ one has $|\gamma_n| \leq 4^n$ with $|\gamma_m| = 4^m$.

(b) From the above observations prove the estimate

$$\left| \frac{f(x_0 + \delta_m) - f(x_0)}{\delta_m} \right| \geq \frac{1}{2}(3^m + 1),$$

and conclude that f is not differentiable at x_0.

5. Let f be a Riemann integrable function on the interval $[-\pi, \pi]$. We define the generalized delayed means of the Fourier series of f by

$$\sigma_{N,K} = \frac{S_N + \cdots + S_{N+K-1}}{K}.$$

Note that in particular

$$\sigma_{0,N} = \sigma_N, \qquad \sigma_{N,1} = S_N \qquad \text{and} \qquad \sigma_{N,N} = \Delta_N,$$

where Δ_N are the specific delayed means used in Section 3.

(a) Show that

$$\sigma_{N,K} = \frac{1}{K}\left((N+K)\sigma_{N+K} - N\sigma_N\right),$$

and

$$\sigma_{N,K} = S_N + \sum_{N+1 \leq |\nu| \leq N+K-1} \left(1 - \frac{|\nu| - N}{K}\right) \hat{f}(\nu) e^{i\nu\theta}.$$

From this last expression for $\sigma_{N,K}$ conclude that

$$|\sigma_{N,K} - S_M| \leq \sum_{N+1 \leq |\nu| \leq N+K-1} |\hat{f}(\nu)|$$

for all $N \leq M < N + K$.

(b) Use one of the above formulas and Fejér's theorem to show that with $N = kn$ and $K = n$, then

$$\sigma_{kn,n}(f)(\theta) \to f(\theta) \qquad \text{as } n \to \infty$$

whenever f is continuous at θ, and also

$$\sigma_{kn,n}(f)(\theta) \to \frac{f(\theta^+) + f(\theta^-)}{2} \qquad \text{as } n \to \infty$$

at a jump discontinuity (refer to the preceding chapters and their exercises for the appropriate definitions and results). In the case when f is continuous on $[-\pi, \pi]$, show that $\sigma_{kn,n}(f) \to f$ uniformly as $n \to \infty$.

(c) Using part (a), show that if $\hat{f}(\nu) = O(1/|\nu|)$ and $kn \leq m < (k+1)n$, we get

$$|\sigma_{kn,n} - S_m| \leq \frac{C}{k} \qquad \text{for some constant } C > 0.$$

(d) Suppose that $\hat{f}(\nu) = O(1/|\nu|)$. Prove that if f is continuous at θ then

$$S_N(f)(\theta) \to f(\theta) \qquad \text{as } N \to \infty,$$

and if f has a jump discontinuity at θ then

$$S_N(f)(\theta) \to \frac{f(\theta^+) + f(\theta^-)}{2} \qquad \text{as } N \to \infty.$$

Also, show that if f is continuous on $[-\pi, \pi]$, then $S_N(f) \to f$ uniformly.

(e) The above arguments show that if $\sum c_n$ is Cesàro summable to s and $c_n = O(1/n)$, then $\sum c_n$ converges to s. This is a weak version of Littlewood's theorem (Problem 3, Chapter 2).

6. Dirichlet's theorem states that the Fourier series of a real continuous periodic function f which has only a finite number of relative maxima and minima converges everywhere to f (and uniformly).

Prove this theorem by showing that such a function satisfies $\hat{f}(n) = O(1/|n|)$.
[Hint: Argue as in Exercise 17, Chapter 3; then use conclusion (d) in Problem 5 above.]

5 The Fourier Transform on $\mathbb{R}$

> The theory of Fourier series and integrals has always had major difficulties and necessitated a large mathematical apparatus in dealing with questions of convergence. It engendered the development of methods of summation, although these did not lead to a completely satisfactory solution of the problem.... For the Fourier transform, the introduction of distributions (hence the space $\mathcal{S}$) is inevitable either in an explicit or hidden form.... As a result one may obtain all that is desired from the point of view of the continuity and inversion of the Fourier transform.
>
> *L. Schwartz,* 1950

The theory of Fourier series applies to functions on the circle, or equivalently, periodic functions on $\mathbb{R}$. In this chapter, we develop an analogous theory for functions on the entire real line which are non-periodic. The functions we consider will be suitably "small" at infinity. There are several ways of defining an appropriate notion of "smallness," but it will nevertheless be vital to assume some sort of vanishing at infinity.

On the one hand, recall that the Fourier series of a periodic function associates a sequence of numbers, namely the Fourier coefficients, to that function; on the other hand, given a suitable function f on $\mathbb{R}$, the analogous object associated to f will in fact be another function $\hat{f}$ on $\mathbb{R}$ which is called the Fourier transform of f. Since the Fourier transform of a function on $\mathbb{R}$ is again a function on $\mathbb{R}$, one can observe a symmetry between a function and its Fourier transform, whose analogue is not as apparent in the setting of Fourier series.

Roughly speaking, the Fourier transform is a continuous version of the Fourier coefficients. Recall that the Fourier coefficients a_n of a function f defined on the circle are given by

$$a_n = \int_0^1 f(x)e^{-2\pi inx}\,dx, \qquad (1)$$

and then in the appropriate sense we have

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x}. \tag{2}$$

Here we have replaced θ by $2\pi x$, as we have frequently done previously.

Now, consider the following analogy where we replace all of the discrete symbols (such as integers and sums) by their continuous counterparts (such as real numbers and integrals). In other words, given a function f on all of $\mathbb{R}$, we define its Fourier transform by changing the domain of integration from the circle to all of $\mathbb{R}$, and by replacing $n \in \mathbb{Z}$ by $\xi \in \mathbb{R}$ in (1), that is, by setting

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi}\, dx. \tag{3}$$

We push our analogy further, and consider the following continuous version of (2): replacing the sum by an integral, and a_n by $\hat{f}(\xi)$, leads to the Fourier inversion formula,

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x \xi}\, d\xi. \tag{4}$$

Under a suitable hypotheses on f, the identity (4) actually holds, and much of the theory in this chapter aims at proving and exploiting this relation. The validity of the Fourier inversion formula is also suggested by the following simple observation. Suppose f is supported in a finite interval contained in $I = [-L/2, L/2]$, and we expand f in a Fourier series on I. Then, letting L tend to infinity, we are led to (4) (see Exercise 1).

The special properties of the Fourier transform make it an important tool in the study of partial differential equations. For instance, we shall see how the Fourier inversion formula allows us to analyze some equations that are modeled on the real line. In particular, following the ideas developed on the circle, we solve the time-dependent heat equation for an infinite rod and the steady-state heat equation in the upper half-plane.

In the last part of the chapter we discuss further topics related to the Poisson summation formula,

$$\sum_{n \in \mathbb{Z}} f(n) = \sum_{n \in \mathbb{Z}} \hat{f}(n),$$

which gives another remarkable connection between periodic functions (and their Fourier series) and non-periodic functions on the line (and

their Fourier transforms). This identity allows us to prove an assertion made in the previous chapter, namely, that the heat kernel $H_t(x)$ satisfies the properties of a good kernel. In addition, the Poisson summation formula arises in many other settings, in particular in parts of number theory, as we shall see in Book II.

We make a final comment about the approach we have chosen. In our study of Fourier series, we found it useful to consider Riemann integrable functions on the circle. In particular, this generality assured us that even functions that had certain discontinuities could be treated by the theory. In contrast, our exposition of the elementary properties of the Fourier transform is stated in terms of the Schwartz space $\mathcal{S}$ of testing functions. These are functions that are indefinitely differentiable and that, together with their derivatives, are rapidly decreasing at infinity. The reliance on this space of functions is a device that allows us to come quickly to the main conclusions, formulated in a direct and transparent fashion. Once this is carried out, we point out some easy extensions to a somewhat wider setting. The more general theory of Fourier transforms (which must necessarily be based on Lebesgue integration) will be treated in Book III.

1 Elementary theory of the Fourier transform

We begin by extending the notion of integration to functions that are defined on the whole real line.

1.1 Integration of functions on the real line

Given the notion of the integral of a function on a closed and bounded interval, the most natural extension of this definition to continuous functions over $\mathbb{R}$ is

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{N\to\infty}\int_{-N}^{N} f(x)\,dx.$$

Of course, this limit may not exist. For example, it is clear that if $f(x)=1$, or even if $f(x)=1/(1+|x|)$, then the above limit is infinite. A moment's reflection suggests that the limit will exist if we impose on f enough decay as $|x|$ tends to infinity. A useful condition is as follows.

A function f defined on $\mathbb{R}$ is said to be of **moderate decrease** if f is continuous and there exists a constant $A>0$ so that

$$|f(x)| \leq \frac{A}{1+x^2} \qquad \text{for all } x\in\mathbb{R}.$$

This inequality says that f is bounded (by A for instance), and also that it decays at infinity at least as fast as $1/x^2$, since $A/(1+x^2) \leq A/x^2$.

For example, the function $f(x) = 1/(1+|x|^n)$ is of moderate decrease as long as $n \geq 2$. Another example is given by the function $e^{-a|x|}$ for $a > 0$.

We shall denote by $\mathcal{M}(\mathbb{R})$ the set of functions of moderate decrease on $\mathbb{R}$. As an exercise, the reader can check that under the usual addition of functions and multiplication by scalars, $\mathcal{M}(\mathbb{R})$ forms a vector space over $\mathbb{C}$.

We next see that whenever f belongs to $\mathcal{M}(\mathbb{R})$, then we may define

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{N\to\infty} \int_{-N}^{N} f(x)\,dx, \tag{5}$$

where the limit now exists. Indeed, for each N the integral $I_N = \int_{-N}^{N} f(x)\,dx$ is well defined because f is continuous. It now suffices to show that $\{I_N\}$ is a Cauchy sequence, and this follows because if $M > N$, then

$$\begin{aligned} |I_M - I_N| &\leq \left| \int_{N\leq|x|\leq M} f(x)\,dx \right| \\ &\leq A \int_{N\leq|x|\leq M} \frac{dx}{x^2} \\ &\leq \frac{2A}{N} \to 0 \qquad \text{as } N \to \infty. \end{aligned}$$

Notice we have also proved that $\int_{|x|\geq N} f(x)\,dx \to 0$ as $N \to \infty$. At this point, we remark that we may replace the exponent 2 in the definition of moderate decrease by $1+\epsilon$ where $\epsilon > 0$; that is,

$$|f(x)| \leq \frac{A}{1+|x|^{1+\epsilon}} \qquad \text{for all } x \in \mathbb{R}.$$

This definition would work just as well for the purpose of the theory developed in this chapter. We chose $\epsilon = 1$ merely as a matter of convenience.

We summarize some elementary properties of integration over $\mathbb{R}$ in a proposition.

Proposition 1.1 *The integral of a function of moderate decrease defined by (5) satisfies the following properties:*

(i) *Linearity: if $f, g \in \mathcal{M}(\mathbb{R})$ and $a, b \in \mathbb{C}$, then*

$$\int_{-\infty}^{\infty} (af(x) + bg(x))\, dx = a \int_{-\infty}^{\infty} f(x)\, dx + b \int_{-\infty}^{\infty} g(x)\, dx.$$

(ii) *Translation invariance: for every $h \in \mathbb{R}$ we have*

$$\int_{-\infty}^{\infty} f(x - h)\, dx = \int_{-\infty}^{\infty} f(x)\, dx.$$

(iii) *Scaling under dilations: if $\delta > 0$, then*

$$\delta \int_{-\infty}^{\infty} f(\delta x)\, dx = \int_{-\infty}^{\infty} f(x)\, dx.$$

(iv) *Continuity: if $f \in \mathcal{M}(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} |f(x - h) - f(x)|\, dx \to 0 \qquad \textit{as } h \to 0.$$

We say a few words about the proof. Property (i) is immediate. To verify property (ii), it suffices to see that

$$\int_{-N}^{N} f(x - h)\, dx - \int_{-N}^{N} f(x)\, dx \to 0 \qquad \text{as } N \to \infty.$$

Since $\int_{-N}^{N} f(x - h)\, dx = \int_{-N-h}^{N-h} f(x)\, dx$, the above difference is majorized by

$$\left| \int_{-N-h}^{-N} f(x)\, dx \right| + \left| \int_{N-h}^{N} f(x)\, dx \right| \leq \frac{A'}{1 + N^2}$$

for large N, which tends to 0 as N tends to infinity.

The proof of property (iii) is similar once we observe that $\delta \int_{-N}^{N} f(\delta x)\, dx = \int_{-\delta N}^{\delta N} f(x)\, dx$.

To prove property (iv) it suffices to take $|h| \leq 1$. For a preassigned $\epsilon > 0$, we first choose N so large that

$$\int_{|x| \geq N} |f(x)|\, dx \leq \epsilon/4 \qquad \text{and} \qquad \int_{|x| \geq N} |f(x - h)|\, dx \leq \epsilon/4.$$

Now with N fixed, we use the fact that since f is continuous, it is uniformly continuous in the interval $[-N - 1, N + 1]$. Hence

$\sup_{|x|\leq N} |f(x-h) - f(x)| \to 0$ as h tends to 0. So we can take h so small that this supremum is less than $\epsilon/4N$. Altogether, then,

$$\begin{aligned}\int_{-\infty}^{\infty} |f(x-h) - f(x)|\, dx &\leq \int_{-N}^{N} |f(x-h) - f(x)|\, dx \\ &\quad + \int_{|x|\geq N} |f(x-h)|\, dx \\ &\quad + \int_{|x|\geq N} |f(x)|\, dx \\ &\leq \epsilon/2 + \epsilon/4 + \epsilon/4 = \epsilon,\end{aligned}$$

and thus conclusion (iv) follows.

1.2 Definition of the Fourier transform

If $f \in \mathcal{M}(\mathbb{R})$, we define its **Fourier transform** for $\xi \in \mathbb{R}$ by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x\xi}\, dx.$$

Of course, $|e^{-2\pi i x\xi}| = 1$, so the integrand is of moderate decrease, and the integral makes sense.

In fact, this last observation implies that $\hat{f}$ is bounded, and moreover, a simple argument shows that $\hat{f}$ is continuous and tends to 0 as $|\xi| \to \infty$ (Exercise 5). However, nothing in the definition above guarantees that $\hat{f}$ is of moderate decrease, or has a specific decay. In particular, it is not clear in this context how to make sense of the integral $\int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x\xi}\, d\xi$ and the resulting Fourier inversion formula. To remedy this, we introduce a more refined space of functions considered by Schwartz which is very useful in establishing the initial properties of the Fourier transform.

The choice of the Schwartz space is motivated by an important principle which ties the decay of $\hat{f}$ to the continuity and differentiability properties of f (and vice versa): the faster $\hat{f}(\xi)$ decreases as $|\xi| \to \infty$, the "smoother" f must be. An example that reflects this principle is given in Exercise 3. We also note that this relationship between f and $\hat{f}$ is reminiscent of a similar one between the smoothness of a function on the circle and the decay of its Fourier coefficients; see the discussion of Corollary 2.4 in Chapter 2.

1.3 The Schwartz space

The **Schwartz space** on $\mathbb{R}$ consists of the set of all indefinitely differentiable functions f so that f and all its derivatives $f', f'', \ldots, f^{(\ell)}, \ldots,$

are **rapidly decreasing**, in the sense that

$$\sup_{x\in\mathbb{R}} |x|^k |f^{(\ell)}(x)| < \infty \qquad \text{for every } k, \ell \geq 0.$$

We denote this space by $\mathcal{S} = \mathcal{S}(\mathbb{R})$, and again, the reader should verify that $\mathcal{S}(\mathbb{R})$ is a vector space over $\mathbb{C}$. Moreover, if $f \in \mathcal{S}(\mathbb{R})$, we have

$$f'(x) = \frac{df}{dx} \in \mathcal{S}(\mathbb{R}) \quad \text{and} \quad xf(x) \in \mathcal{S}(\mathbb{R}).$$

This expresses the important fact that the Schwartz space is closed under differentiation and multiplication by polynomials.

A simple example of a function in $\mathcal{S}(\mathbb{R})$ is the **Gaussian** defined by

$$f(x) = e^{-x^2},$$

which plays a central role in the theory of the Fourier transform, as well as other fields (for example, probability theory and physics). The reader can check that the derivatives of f are of the form $P(x)e^{-x^2}$ where P is a polynomial, and this immediately shows that $f \in \mathcal{S}(\mathbb{R})$. In fact, e^{-ax^2} belongs to $\mathcal{S}(\mathbb{R})$ whenever $a > 0$. Later, we will normalize the Gaussian by choosing $a = \pi$.

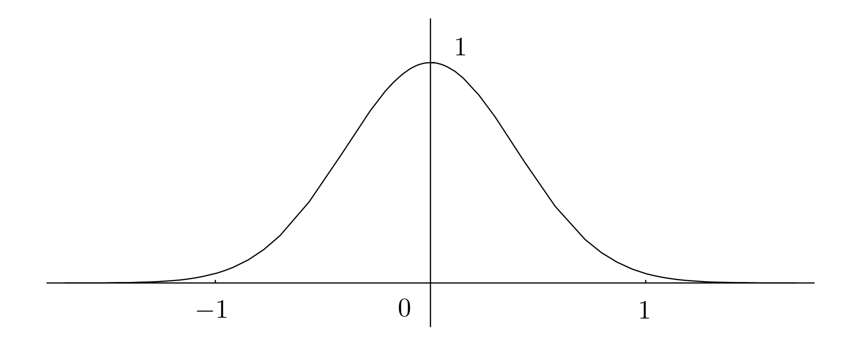


Figure 1. The Gaussian $e^{-\pi x^2}$

An important class of other examples in $\mathcal{S}(\mathbb{R})$ are the "bump functions" which vanish outside bounded intervals (Exercise 4).

As a final remark, note that although $e^{-|x|}$ decreases rapidly at infinity, it is not differentiable at 0 and therefore does not belong to $\mathcal{S}(\mathbb{R})$.

1.4 The Fourier transform on $\mathcal{S}$

The **Fourier transform** of a function $f \in \mathcal{S}(\mathbb{R})$ is defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ix\xi}\, dx.$$

Some simple properties of the Fourier transform are gathered in the following proposition. We use the notation

$$f(x) \longrightarrow \hat{f}(\xi)$$

to mean that $\hat{f}$ denotes the Fourier transform of f.

Proposition 1.2 *If $f \in \mathcal{S}(\mathbb{R})$ then:*

(i) $f(x+h) \longrightarrow \hat{f}(\xi)e^{2\pi ih\xi}$ *whenever* $h \in \mathbb{R}$.

(ii) $f(x)e^{-2\pi ixh} \longrightarrow \hat{f}(\xi + h)$ *whenever* $h \in \mathbb{R}$.

(iii) $f(\delta x) \longrightarrow \delta^{-1}\hat{f}(\delta^{-1}\xi)$ *whenever* $\delta > 0$.

(iv) $f'(x) \longrightarrow 2\pi i\xi\hat{f}(\xi)$.

(v) $-2\pi ixf(x) \longrightarrow \dfrac{d}{d\xi}\hat{f}(\xi)$.

In particular, except for factors of $2\pi i$, the Fourier transform interchanges differentiation and multiplication by x. This is the key property that makes the Fourier transform a central object in the theory of differential equations. We shall return to this point later.

Proof. Property (i) is an immediate consequence of the translation invariance of the integral, and property (ii) follows from the definition. Also, the third property of Proposition 1.1 establishes (iii).

Integrating by parts gives

$$\int_{-N}^{N} f'(x)e^{-2\pi ix\xi}\, dx = \left[f(x)e^{-2\pi ix\xi}\right]_{-N}^{N} + 2\pi i\xi \int_{-N}^{N} f(x)e^{-2\pi ix\xi}\, dx,$$

so letting N tend to infinity gives (iv).

Finally, to prove property (v), we must show that $\hat{f}$ is differentiable and find its derivative. Let $\epsilon > 0$ and consider

$$\frac{\hat{f}(\xi+h) - \hat{f}(\xi)}{h} - (\widehat{-2\pi ixf})(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ix\xi}\left[\frac{e^{-2\pi ixh} - 1}{h} + 2\pi ix\right] dx.$$

Since $f(x)$ and $xf(x)$ are of rapid decrease, there exists an integer N so that $\int_{|x|\geq N} |f(x)|\,dx \leq \epsilon$ and $\int_{|x|\geq N} |x|\,|f(x)|\,dx \leq \epsilon$. Moreover, for $|x| \leq N$, there exists h_0 so that $|h| < h_0$ implies

$$\left|\frac{e^{-2\pi ixh}-1}{h} + 2\pi ix\right| \leq \frac{\epsilon}{N}.$$

Hence for $|h| < h_0$ we have

$$\begin{aligned}\left|\frac{\hat{f}(\xi+h)-\hat{f}(\xi)}{h} - (\widehat{-2\pi ixf})(\xi)\right| & \\ &\leq \int_{-N}^{N}\left|f(x)e^{-2\pi ix\xi}\left[\frac{e^{-2\pi ixh}-1}{h}+2\pi ix\right]\right| dx + C\epsilon \\ &\leq C'\epsilon.\end{aligned}$$

Theorem 1.3 *If $f \in \mathcal{S}(\mathbb{R})$, then $\hat{f} \in \mathcal{S}(\mathbb{R})$.*

The proof is an easy application of the fact that the Fourier transform interchanges differentiation and multiplication. In fact, note that if $f \in \mathcal{S}(\mathbb{R})$, its Fourier transform $\hat{f}$ is bounded; then also, for each pair of non-negative integers ℓ and k, the expression

$$\xi^k \left(\frac{d}{d\xi}\right)^{\ell} \hat{f}(\xi)$$

is bounded, since by the last proposition, it is the Fourier transform of

$$\frac{1}{(2\pi i)^k}\left(\frac{d}{dx}\right)^k [(-2\pi ix)^{\ell} f(x)].$$

The proof of the inversion formula

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi ix\xi}\,d\xi \quad \text{for } f \in \mathcal{S}(\mathbb{R}),$$

which we give in the next section, is based on a careful study of the function e^{-ax^2}, which, as we have already observed, is in $\mathcal{S}(\mathbb{R})$ if $a > 0$.

The Gaussians as good kernels

We begin by considering the case $a = \pi$ because of the normalization:

$$\int_{-\infty}^{\infty} e^{-\pi x^2}\, dx = 1. \tag{6}$$

To see why (6) is true, we use the multiplicative property of the exponential to reduce the calculation to a two-dimensional integral. More precisely, we can argue as follows:

$$\begin{aligned}
\left(\int_{-\infty}^{\infty} e^{-\pi x^2}\, dx\right)^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\pi(x^2+y^2)}\, dx\, dy \\
&= \int_0^{2\pi}\int_0^{\infty} e^{-\pi r^2} r\, dr\, d\theta \\
&= \int_0^{\infty} 2\pi r e^{-\pi r^2}\, dr \\
&= \left[-e^{-\pi r^2}\right]_0^{\infty} \\
&= 1,
\end{aligned}$$

where we have evaluated the two-dimensional integral using polar coordinates.

The fundamental property of the Gaussian which is of interest to us, and which actually follows from (6), is that $e^{-\pi x^2}$ equals its Fourier transform! We isolate this important result in a theorem.

Theorem 1.4 *If $f(x) = e^{-\pi x^2}$, then $\hat{f}(\xi) = f(\xi)$.*

Proof. Define

$$F(\xi) = \hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x \xi}\, dx,$$

and observe that $F(0) = 1$, by our previous calculation. By property (v) in Proposition 1.2, and the fact that $f'(x) = -2\pi x f(x)$, we obtain

$$F'(\xi) = \int_{-\infty}^{\infty} f(x)(-2\pi i x) e^{-2\pi i x \xi}\, dx = i \int_{-\infty}^{\infty} f'(x) e^{-2\pi i x \xi}\, dx.$$

By (iv) of the same proposition, we find that

$$F'(\xi) = i(2\pi i \xi)\hat{f}(\xi) = -2\pi \xi F(\xi).$$

If we define $G(\xi) = F(\xi)e^{\pi\xi^2}$, then from what we have seen above, it follows that $G'(\xi) = 0$, hence G is constant. Since $F(0) = 1$, we conclude that G is identically equal to 1, therefore $F(\xi) = e^{-\pi\xi^2}$, as was to be shown.

The scaling properties of the Fourier transform under dilations yield the following important transformation law, which follows from (iii) in Proposition 1.2 (with δ replaced by $\delta^{-1/2}$).

Corollary 1.5 *If $\delta > 0$ and $K_\delta(x) = \delta^{-1/2}e^{-\pi x^2/\delta}$, then $\widehat{K_\delta}(\xi) = e^{-\pi\delta\xi^2}$.*

We pause to make an important observation. As δ tends to 0, the function K_δ peaks at the origin, while its Fourier transform $\widehat{K_\delta}$ gets flatter. So in this particular example, we see that K_δ and $\widehat{K_\delta}$ cannot both be localized (that is, concentrated) at the origin. This is an example of a general phenomenon called the Heisenberg uncertainty principle, which we will discuss at the end of this chapter.

We have now constructed a family of good kernels on the real line, analogous to those on the circle considered in Chapter 2. Indeed, with

$$K_\delta(x) = \delta^{-1/2}e^{-\pi x^2/\delta},$$

we have:

(i) $\int_{-\infty}^{\infty} K_\delta(x)\,dx = 1$.

(ii) $\int_{-\infty}^{\infty} |K_\delta(x)|\,dx \le M$.

(iii) For every $\eta > 0$, we have $\int_{|x|>\eta} |K_\delta(x)|\,dx \to 0$ as $\delta \to 0$.

To prove (i), we may change variables and use (6), or note that the integral equals $\widehat{K_\delta}(0)$, which is 1 by Corollary 1.5. Since $K_\delta \ge 0$, it is clear that property (ii) is also true. Finally we can again change variables to get

$$\int_{|x|>\eta} |K_\delta(x)|\,dx = \int_{|y|>\eta/\delta^{1/2}} e^{-\pi y^2}\,dy \to 0$$

as δ tends to 0. We have thus proved the following result.

Theorem 1.6 *The collection $\{K_\delta\}_{\delta>0}$ is a family of good kernels as $\delta \to 0$.*

We next apply these good kernels via the operation of convolution, which is given as follows. If $f, g \in \mathcal{S}(\mathbb{R})$, their **convolution** is defined by

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-t)g(t)\,dt. \tag{7}$$

For a fixed value of x, the function $f(x-t)g(t)$ is of rapid decrease in t, hence the integral converges.

By the argument in Section 4 of Chapter 2 (with a slight modification), we get the following corollary.

Corollary 1.7 *If $f \in \mathcal{S}(\mathbb{R})$, then*

$$(f * K_\delta)(x) \to f(x) \quad \text{uniformly in } x \text{ as } \delta \to 0.$$

Proof. First, we claim that f is uniformly continuous on $\mathbb{R}$. Indeed, given $\epsilon > 0$ there exists $R > 0$ so that $|f(x)| < \epsilon/4$ whenever $|x| \geq R$. Moreover, f is continuous, hence uniformly continuous on the compact interval $[-R, R]$, and together with the previous observation, we can find $\eta > 0$ so that $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \eta$. Now we argue as usual. Using the first property of good kernels, we can write

$$(f * K_\delta)(x) - f(x) = \int_{-\infty}^{\infty} K_\delta(t)\,[f(x-t) - f(x)]\,dt,$$

and since $K_\delta \geq 0$, we find

$$|(f * K_\delta)(x) - f(x)| \leq \int_{|t|>\eta} + \int_{|t|\leq\eta} K_\delta(t)\,|f(x-t) - f(x)|\,dt.$$

The first integral is small by the third property of good kernels, and the fact that f is bounded, while the second integral is also small since f is uniformly continuous and $\int K_\delta = 1$. This concludes the proof of the corollary.

1.5 The Fourier inversion

The next result is an identity sometimes called the multiplication formula.

Proposition 1.8 *If $f, g \in \mathcal{S}(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} f(x)\hat{g}(x)\,dx = \int_{-\infty}^{\infty} \hat{f}(y)g(y)\,dy.$$

To prove the proposition, we need to digress briefly to discuss the interchange of the order of integration for double integrals. Suppose $F(x, y)$ is a continuous function in the plane $(x, y) \in \mathbb{R}^2$. We will assume the following decay condition on F:

$$|F(x, y)| \leq A/(1 + x^2)(1 + y^2).$$

Then, we can state that for each x the function $F(x, y)$ is of moderate decrease in y, and similarly for each fixed y the function $F(x, y)$ is of moderate decrease in x. Moreover, the function $F_1(x) = \int_{-\infty}^{\infty} F(x, y)\, dy$ is continuous and of moderate decrease; similarly for the function $F_2(y) = \int_{-\infty}^{\infty} F(x, y)\, dx$. Finally

$$\int_{-\infty}^{\infty} F_1(x)\, dx = \int_{-\infty}^{\infty} F_2(y)\, dy.$$

The proof of these facts may be found in the appendix.

We now apply this to $F(x, y) = f(x)g(y)e^{-2\pi ixy}$. Then $F_1(x) = f(x)\hat{g}(x)$, and $F_2(y) = \hat{f}(y)g(y)$ so

$$\int_{-\infty}^{\infty} f(x)\hat{g}(x)\, dx = \int_{-\infty}^{\infty} \hat{f}(y)g(y)\, dy,$$

which is the assertion of the proposition.

The multiplication formula and the fact that the Gaussian is its own Fourier transform lead to a proof of the first major theorem.

Theorem 1.9 (Fourier inversion) *If $f \in \mathcal{S}(\mathbb{R})$, then*

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi ix\xi}\, d\xi.$$

Proof. We first claim that

$$f(0) = \int_{-\infty}^{\infty} \hat{f}(\xi)\, d\xi.$$

Let $G_\delta(x) = e^{-\pi\delta x^2}$ so that $\widehat{G_\delta}(\xi) = K_\delta(\xi)$. By the multiplication formula we get

$$\int_{-\infty}^{\infty} f(x)K_\delta(x)\, dx = \int_{-\infty}^{\infty} \hat{f}(\xi)G_\delta(\xi)\, d\xi.$$

Since K_δ is a good kernel, the first integral goes to $f(0)$ as δ tends to 0. Since the second integral clearly converges to $\int_{-\infty}^{\infty} \hat{f}(\xi)\, d\xi$ as δ tends to 0, our claim is proved. In general, let $F(y) = f(y + x)$ so that

$$f(x) = F(0) = \int_{-\infty}^{\infty} \hat{F}(\xi)\, d\xi = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi ix\xi}\, d\xi.$$

As the name of Theorem 1.9 suggests, it provides a formula that inverts the Fourier transform; in fact we see that the Fourier transform is its own

inverse except for the change of x to $-x$. More precisely, we may define two mappings $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ and $\mathcal{F}^* : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ by

$$\mathcal{F}(f)(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ix\xi}\, dx \quad \text{and} \quad \mathcal{F}^*(g)(x) = \int_{-\infty}^{\infty} g(\xi)e^{2\pi ix\xi}\, d\xi.$$

Thus $\mathcal{F}$ is the Fourier transform, and Theorem 1.9 guarantees that $\mathcal{F}^* \circ \mathcal{F} = I$ on $\mathcal{S}(\mathbb{R})$, where I is the identity mapping. Moreover, since the definitions of $\mathcal{F}$ and $\mathcal{F}^*$ differ only by a sign in the exponential, we see that $\mathcal{F}(f)(y) = \mathcal{F}^*(f)(-y)$, so we also have $\mathcal{F} \circ \mathcal{F}^* = I$. As a consequence, we conclude that $\mathcal{F}^*$ is the inverse of the Fourier transform on $\mathcal{S}(\mathbb{R})$, and we get the following result.

Corollary 1.10 *The Fourier transform is a bijective mapping on the Schwartz space.*

1.6 The Plancherel formula

We need a few further results about convolutions of Schwartz functions. The key fact is that the Fourier transform interchanges convolutions with pointwise products, a result analogous to the situation for Fourier series.

Proposition 1.11 *If $f, g \in \mathcal{S}(\mathbb{R})$ then:*

(i) $f * g \in \mathcal{S}(\mathbb{R})$.

(ii) $f * g = g * f$.

(iii) $\widehat{(f * g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.

Proof. To prove that $f * g$ is rapidly decreasing, observe first that for any $\ell \geq 0$ we have $\sup_x |x|^\ell |g(x-y)| \leq A_\ell (1 + |y|)^\ell$, because g is rapidly decreasing (to check this assertion, consider separately the two cases $|x| \leq 2|y|$ and $|x| \geq 2|y|$). From this, we see that

$$\sup_x |x^\ell (f * g)(x)| \leq A_\ell \int_{-\infty}^{\infty} |f(y)|(1 + |y|)^\ell\, dy,$$

so that $x^\ell (f * g)(x)$ is a bounded function for every $\ell \geq 0$. These estimates carry over to the derivatives of $f * g$, thereby proving that $f * g \in \mathcal{S}(\mathbb{R})$ because

$$\left(\frac{d}{dx}\right)^k (f * g)(x) = (f * \left(\frac{d}{dx}\right)^k g)(x) \quad \text{for } k = 1, 2, \ldots.$$

This identity is proved first for $k = 1$ by differentiating under the integral defining $f * g$. The interchange of differentiation and integration is justified in this case by the rapid decrease of dg/dx. The identity then follows for every k by iteration.

For fixed x, the change of variables $x - y = u$ shows that

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-u)g(u)\,du = (g * f)(x).$$

This change of variables is a composition of two changes, $y \mapsto -y$ and $y \mapsto y - h$ (with $h = x$). For the first one we use the observation that $\int_{-\infty}^{\infty} F(x)\,dx = \int_{-\infty}^{\infty} F(-x)\,dx$ for any Schwartz function F, and for the second, we apply (ii) of Proposition 1.1

Finally, consider $F(x, y) = f(y)g(x - y)e^{-2\pi ix\xi}$. Since f and g are rapidly decreasing, considering separately the two cases $|x| \leq 2|y|$ and $|x| \geq 2|y|$, we see that the discussion of the change of order of integration after Proposition 1.8 applies to F. In this case $F_1(x) = (f * g)(x)e^{-2\pi ix\xi}$, and $F_2(y) = f(y)e^{-2\pi iy\xi}\hat{g}(\xi)$. Thus $\int_{-\infty}^{\infty} F_1(x)\,dx = \int_{-\infty}^{\infty} F_2(y)\,dy$, which implies (iii). The proposition is therefore proved.

We now use the properties of convolutions of Schwartz functions to prove the main result of this section. The result we have in mind is the analogue for functions on $\mathbb{R}$ of Parseval's identity for Fourier series.

The Schwartz space can be equipped with a Hermitian inner product

$$(f, g) = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\,dx$$

whose associated norm is

$$\|f\| = \left(\int_{-\infty}^{\infty} |f(x)|^2\,dx\right)^{1/2}.$$

The second major theorem in the theory states that the Fourier transform is a unitary transformation on $\mathcal{S}(\mathbb{R})$.

Theorem 1.12 (Plancherel) *If $f \in \mathcal{S}(\mathbb{R})$ then $\|\hat{f}\| = \|f\|$.*

Proof. If $f \in \mathcal{S}(\mathbb{R})$ define $f^\flat(x) = \overline{f(-x)}$. Then $\widehat{f^\flat}(\xi) = \overline{\hat{f}(\xi)}$. Now let $h = f * f^\flat$. Clearly, we have

$$\hat{h}(\xi) = |\hat{f}(\xi)|^2 \quad \text{and} \quad h(0) = \int_{-\infty}^{\infty} |f(x)|^2\,dx.$$

The theorem now follows from the inversion formula applied with $x = 0$, that is,

$$\int_{-\infty}^{\infty} \hat{h}(\xi)\, d\xi = h(0).$$

1.7 Extension to functions of moderate decrease

In the previous sections, we have limited our assertion of the Fourier inversion and Plancherel formulas to the case when the function involved belonged to the Schwartz space. It does not really involve further ideas to extend these results to functions of moderate decrease, once we make the additional assumption that the Fourier transform of the function under consideration is also of moderate decrease. Indeed, the key observation, which is easy to prove, is that the convolution $f * g$ of two functions f and g of moderate decrease is again a function of moderate decrease (Exercise 7); also $\widehat{f * g} = \hat{f}\hat{g}$. Moreover, the multiplication formula continues to hold, and we deduce the Fourier inversion and Plancherel formulas when f and $\hat{f}$ are both of moderate decrease.

This generalization, although modest in scope, is nevertheless useful in some circumstances.

1.8 The Weierstrass approximation theorem

We now digress briefly by further exploiting our good kernels to prove the Weierstrass approximation theorem. This result was already alluded to in Chapter 2.

Theorem 1.13 *Let f be a continuous function on the closed and bounded interval $[a, b] \subset \mathbb{R}$. Then, for any $\epsilon > 0$, there exists a polynomial P such that*

$$\sup_{x \in [a,b]} |f(x) - P(x)| < \epsilon.$$

In other words, f can be uniformly approximated by polynomials.

Proof. Let $[-M, M]$ denote any interval that contains $[a, b]$ in its interior, and let g be a continuous function on $\mathbb{R}$ that equals 0 outside $[-M, M]$ and equals f in $[a, b]$. For example, extend f as follows: from b to M define g by a straight line segment going from $f(b)$ to 0, and from a to $-M$ by a straight line segment from $f(a)$ also to 0. Let B be a

bound for g, that is, $|g(x)| \leq B$ for all x. Then, since $\{K_\delta\}$ is a family of good kernels, and g is continuous with compact support, we may argue as in the proof of Corollary 1.7 to see that $g * K_\delta$ converges uniformly to g as δ tends to 0. In fact, we choose δ_0 so that

$$|g(x) - (g * K_{\delta_0})(x)| < \epsilon/2 \quad \text{for all } x \in \mathbb{R}.$$

Now, we recall that e^x is given by the power series expansion $e^x = \sum_{n=0}^{\infty} x^n/n!$ which converges uniformly in every compact interval of $\mathbb{R}$. Therefore, there exists an integer N so that

$$|K_{\delta_0}(x) - R(x)| \leq \frac{\epsilon}{4MB} \quad \text{for all } x \in [-2M, 2M]$$

where $R(x) = \delta_0^{-1/2} \sum_{n=0}^{N} \frac{(-\pi x^2/\delta_0)^n}{n!}$. Then, recalling that g vanishes outside the interval $[-M, M]$, we have that for all $x \in [-M, M]$

$$\begin{aligned} |(g * K_{\delta_0})(x) - (g * R)(x)| &= \left| \int_{-M}^{M} g(t) \left[K_{\delta_0}(x-t) - R(x-t) \right] dt \right| \\ &\leq \int_{-M}^{M} |g(t)| \, |K_{\delta_0}(x-t) - R(x-t)| \, dt \\ &\leq 2MB \sup_{z \in [-2M, 2M]} |K_{\delta_0}(z) - R(z)| \\ &< \epsilon/2. \end{aligned}$$

Therefore, the triangle inequality implies that $|g(x) - (g * R)(x)| < \epsilon$ whenever $x \in [-M, M]$, hence $|f(x) - (g * R)(x)| < \epsilon$ when $x \in [a, b]$.

Finally, note that $g * R$ is a polynomial in the x variable. Indeed, by definition we have $(g * R)(x) = \int_{-M}^{M} g(t) R(x-t) \, dt$, and $R(x-t)$ is a polynomial in x since it can be expressed, after several expansions, as $R(x-t) = \sum_n a_n(t) x^n$ where the sum is finite. This concludes the proof of the theorem.

2 Applications to some partial differential equations

We mentioned earlier that a crucial property of the Fourier transform is that it interchanges differentiation and multiplication by polynomials. We now use this crucial fact together with the Fourier inversion theorem to solve some specific partial differential equations.

2.1 The time-dependent heat equation on the real line

In Chapter 4 we considered the heat equation on the circle. Here we study the analogous problem on the real line.

Consider an infinite rod, which we model by the real line, and suppose that we are given an initial temperature distribution $f(x)$ on the rod at time $t = 0$. We wish now to determine the temperature $u(x, t)$ at a point x at time $t > 0$. Considerations similar to the ones given in Chapter 1 show that when u is appropriately normalized, it solves the following partial differential equation:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \tag{8}$$

called the **heat equation**. The initial condition we impose is $u(x, 0) = f(x)$.

Just as in the case of the circle, the solution is given in terms of a convolution. Indeed, define the **heat kernel** of the line by

$$\mathcal{H}_t(x) = K_\delta(x), \quad \text{with } \delta = 4\pi t,$$

so that

$$\mathcal{H}_t(x) = \frac{1}{(4\pi t)^{1/2}} e^{-x^2/4t} \quad \text{and} \quad \hat{\mathcal{H}}_t(\xi) = e^{-4\pi^2 t\xi^2}.$$

Taking the Fourier transform of equation (8) in the x variable (formally) leads to

$$\frac{\partial \hat{u}}{\partial t}(\xi, t) = -4\pi^2\xi^2\hat{u}(\xi, t).$$

Fixing ξ, this is an ordinary differential equation in the variable t (with unknown $\hat{u}(\xi, \cdot)$), so there exists a constant $A(\xi)$ so that

$$\hat{u}(\xi, t) = A(\xi)e^{-4\pi^2\xi^2 t}.$$

We may also take the Fourier transform of the initial condition and obtain $\hat{u}(\xi, 0) = \hat{f}(\xi)$, hence $A(\xi) = \hat{f}(\xi)$. This leads to the following theorem.

Theorem 2.1 *Given $f \in \mathcal{S}(\mathbb{R})$, let*

$$u(x, t) = (f * \mathcal{H}_t)(x) \quad \textit{for } t > 0$$

where $\mathcal{H}_t$ is the heat kernel. Then:

(i) *The function u is C^2 when $x \in \mathbb{R}$ and $t > 0$, and u solves the heat equation.*

(ii) $u(x,t) \to f(x)$ *uniformly in* x *as* $t \to 0$. *Hence if we set* $u(x,0) = f(x)$, *then* u *is continuous on the closure of the upper half-plane* $\overline{\mathbb{R}^2_+} = \{(x,t) : x \in \mathbb{R},\ t \geq 0\}$.

(iii) $\int_{-\infty}^{\infty} |u(x,t) - f(x)|^2\, dx \to 0$ *as* $t \to 0$.

Proof. Because $u = f * \mathcal{H}_t$, taking the Fourier transform in the x-variable gives $\hat{u} = \hat{f}\hat{\mathcal{H}}_t$, and so $\hat{u}(\xi,t) = \hat{f}(\xi)e^{-4\pi^2\xi^2 t}$. The Fourier inversion formula gives

$$u(x,t) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{-4\pi^2 t\xi^2} e^{2\pi i\xi x}\, d\xi.$$

By differentiating under the integral sign, one verifies (i). In fact, one observes that u is indefinitely differentiable. Note that (ii) is an immediate consequence of Corollary 1.7. Finally, by Plancherel's formula, we have

$$\begin{aligned}\int_{-\infty}^{\infty} |u(x,t) - f(x)|^2\, dx &= \int_{-\infty}^{\infty} |\hat{u}(\xi,t) - \hat{f}(\xi)|^2\, d\xi \\ &= \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2\, |e^{-4\pi^2 t\xi^2} - 1|\, d\xi.\end{aligned}$$

To see that this last integral goes to 0 as $t \to 0$, we argue as follows: since $|e^{-4\pi^2 t\xi^2} - 1| \leq 2$ and $f \in \mathcal{S}(\mathbb{R})$, we can find N so that

$$\int_{|\xi| \geq N} |\hat{f}(\xi)|^2 |e^{-4\pi^2 t\xi^2} - 1|\, d\xi < \epsilon,$$

and for all small t we have $\sup_{|\xi| \leq N} |\hat{f}(\xi)|^2 |e^{-4\pi^2 t\xi^2} - 1| < \epsilon/2N$ since $\hat{f}$ is bounded. Thus

$$\int_{|\xi| \leq N} |\hat{f}(\xi)|^2\, |e^{-4\pi^2 t\xi^2} - 1|\, d\xi < \epsilon \quad \text{for all small } t.$$

This completes the proof of the theorem.

The above theorem guarantees the existence of a solution to the heat equation with initial data f. This solution is also unique, if uniqueness is formulated appropriately. In this regard, we note that $u = f * \mathcal{H}_t$, $f \in \mathcal{S}(\mathbb{R})$, satisfies the following additional property.

Corollary 2.2 $u(\cdot,t)$ *belongs to* $\mathcal{S}(\mathbb{R})$ *uniformly in* t, *in the sense that for any* $T > 0$

$$\sup_{\substack{x \in \mathbb{R} \\ 0 < t < T}} |x|^k \left| \frac{\partial^\ell}{\partial x^\ell} u(x,t) \right| < \infty \quad \textit{for each } k, \ell \geq 0. \tag{9}$$

Proof. This result is a consequence of the following estimate:

$$|u(x,t)| \leq \int_{|y|\leq|x|/2} |f(x-y)|\mathcal{H}_t(y)\,dy + \int_{|y|\geq|x|/2} |f(x-y)|\mathcal{H}_t(y)\,dy$$
$$\leq \frac{C_N}{(1+|x|)^N} + \frac{C}{\sqrt{t}}e^{-cx^2/t}.$$

Indeed, since f is rapidly decreasing, we have $|f(x-y)| \leq C_N/(1+|x|)^N$ when $|y| \leq |x|/2$. Also, if $|y| \geq |x|/2$ then $\mathcal{H}_t(y) \leq Ct^{-1/2}e^{-cx^2/t}$, and we obtain the above inequality. Consequently, we see that $u(x,t)$ is rapidly decreasing uniformly for $0 < t < T$.

The same argument can be applied to the derivatives of u in the x variable since we may differentiate under the integral sign and apply the above estimate with f replaced by f', and so on.

This leads to the following uniqueness theorem.

Theorem 2.3 *Suppose $u(x,t)$ satisfies the following conditions:*

(i) *u is continuous on the closure of the upper half-plane.*

(ii) *u satisfies the heat equation for $t > 0$.*

(iii) *u satisfies the boundary condition $u(x,0) = 0$.*

(iv) *$u(\cdot,t) \in \mathcal{S}(\mathbb{R})$ uniformly in t, as in (9).*

Then, we conclude that $u = 0$.

Below we use the abbreviations $\partial_x^\ell u$ and $\partial_t u$ to denote $\partial^\ell u/\partial x^\ell$ and $\partial u/\partial t$, respectively.

Proof. We define the energy at time t of the solution $u(x,t)$ by

$$E(t) = \int_{\mathbb{R}} |u(x,t)|^2\,dx.$$

Clearly $E(t) \geq 0$. Since $E(0) = 0$ it suffices to show that E is a decreasing function, and this is achieved by proving that $dE/dt \leq 0$. The assumptions on u allow us to differentiate $E(t)$ under the integral sign

$$\frac{dE}{dt} = \int_{\mathbb{R}} [\partial_t u(x,t)\overline{u}(x,t) + u(x,t)\partial_t\overline{u}(x,t)]\,dx.$$

But u satisfies the heat equation, therefore $\partial_t u = \partial_x^2 u$ and $\partial_t \overline{u} = \partial_x^2\overline{u}$, so that after an integration by parts, where we use the fact that u and its

x derivatives decrease rapidly as $|x| \to \infty$, we find

$$\begin{aligned}\frac{dE}{dt} &= \int_{\mathbb{R}} \left[\partial_x^2 u(x,t)\overline{u}(x,t) + u(x,t)\partial_x^2\overline{u}(x,t)\right]\, dx \\ &= -\int_{\mathbb{R}} \left[\partial_x u(x,t)\partial_x\overline{u}(x,t) + \partial_x u(x,t)\partial_x\overline{u}(x,t)\right]\, dx \\ &= -2\int_{\mathbb{R}} |\partial_x u(x,t)|^2\, dx \\ &\leq 0,\end{aligned}$$

as claimed. Thus $E(t) = 0$ for all t, hence $u = 0$.

Another uniqueness theorem for the heat equation, with a less restrictive assumption than (9), can be found in Problem 6. Examples when uniqueness fails are given in Exercise 12 and Problem 4.

2.2 The steady-state heat equation in the upper half-plane

The equation we are now concerned with is

$$\triangle u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{10}$$

in the upper half-plane $\mathbb{R}^2_+ = \{(x,y) : x \in \mathbb{R},\ y > 0\}$. The boundary condition we require is $u(x,0) = f(x)$. The operator $\triangle$ is the Laplacian and the above partial differential equation describes the steady-state heat distribution in $\mathbb{R}^2_+$ subject to $u = f$ on the boundary. The kernel that solves this problem is called the **Poisson kernel** for the upper half-plane, and is given by

$$\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2} \qquad \text{where } x \in \mathbb{R} \text{ and } y > 0.$$

This is the analogue of the Poisson kernel for the disc discussed in Section 5.4 of Chapter 2.

Note that for each fixed y the kernel $\mathcal{P}_y$ is only of moderate decrease as a function of x, so we will use the theory of the Fourier transform appropriate for these types of functions (see Section 1.7).

We proceed as in the case of the time-dependent heat equation, by taking the Fourier transform of equation (10) (formally) in the x variable, thereby obtaining

$$-4\pi^2\xi^2\hat{u}(\xi,y) + \frac{\partial^2\hat{u}}{\partial y^2}(\xi,y) = 0$$

with the boundary condition $\hat{u}(\xi, 0) = \hat{f}(\xi)$. The general solution of this ordinary differential equation in y (with ξ fixed) takes the form

$$\hat{u}(\xi, y) = A(\xi)e^{-2\pi|\xi|y} + B(\xi)e^{2\pi|\xi|y}.$$

If we disregard the second term because of its rapid exponential increase we find, after setting $y = 0$, that

$$\hat{u}(\xi, y) = \hat{f}(\xi)e^{-2\pi|\xi|y}.$$

Therefore u is given in terms of the convolution of f with a kernel whose Fourier transform is $e^{-2\pi|\xi|y}$. This is precisely the Poisson kernel given above, as we prove next.

Lemma 2.4 *The following two identities hold:*

$$\int_{-\infty}^{\infty} e^{-2\pi|\xi|y}e^{2\pi i\xi x}\,d\xi = \mathcal{P}_y(x),$$
$$\int_{-\infty}^{\infty} \mathcal{P}_y(x)e^{-2\pi ix\xi}\,dx = e^{-2\pi|\xi|y}.$$

Proof. The first formula is fairly straightforward since we can split the integral from $-\infty$ to 0 and 0 to ∞. Then, since $y > 0$ we have

$$\int_0^{\infty} e^{-2\pi\xi y}e^{2\pi i\xi x}\,d\xi = \int_0^{\infty} e^{2\pi i(x+iy)\xi}\,d\xi = \left[\frac{e^{2\pi i(x+iy)\xi}}{2\pi i(x+iy)}\right]_0^{\infty} = -\frac{1}{2\pi i(x+iy)},$$

and similarly,

$$\int_{-\infty}^{0} e^{2\pi\xi y}e^{2\pi i\xi x}\,d\xi = \frac{1}{2\pi i(x-iy)}.$$

Therefore

$$\int_{-\infty}^{\infty} e^{-2\pi|\xi|y}e^{2\pi i\xi x}\,d\xi = \frac{1}{2\pi i(x-iy)} - \frac{1}{2\pi i(x+iy)} = \frac{y}{\pi(x^2+y^2)}.$$

The second formula is now a consequence of the Fourier inversion theorem applied in the case when f and $\hat{f}$ are of moderate decrease.

Lemma 2.5 *The Poisson kernel is a good kernel on* $\mathbb{R}$ *as* $y \to 0$.

Proof. Setting $\xi = 0$ in the second formula of the lemma shows that $\int_{-\infty}^{\infty} \mathcal{P}_y(x)\,dx = 1$, and clearly $\mathcal{P}_y(x) \geq 0$, so it remains to check the last property of good kernels. Given a fixed $\delta > 0$, we may change variables $u = x/y$ so that

$$\int_{\delta}^{\infty} \frac{y}{x^2+y^2}\,dx = \int_{\delta/y}^{\infty} \frac{du}{1+u^2} = [\arctan u]_{\delta/y}^{\infty} = \pi/2 - \arctan(\delta/y),$$

and this quantity goes to 0 as $y \to 0$. Since $\mathcal{P}_y(x)$ is an even function, the proof is complete.

The following theorem establishes the existence of a solution to our problem.

Theorem 2.6 *Given $f \in \mathcal{S}(\mathbb{R})$, let $u(x,y) = (f * \mathcal{P}_y)(x)$. Then:*

(i) *$u(x,y)$ is C^2 in $\mathbb{R}^2_+$ and $\triangle u = 0$.*

(ii) *$u(x,y) \to f(x)$ uniformly as $y \to 0$.*

(iii) *$\int_{-\infty}^{\infty} |u(x,y) - f(x)|^2\,dx \to 0$ as $y \to 0$.*

(iv) *If $u(x,0) = f(x)$, then u is continuous on the closure $\overline{\mathbb{R}^2_+}$ of the upper half-plane, and vanishes at infinity in the sense that*

$$u(x,y) \to 0 \qquad \text{as } |x| + y \to \infty.$$

Proof. The proofs of parts (i), (ii), and (iii) are similar to the case of the heat equation, and so are left to the reader. Part (iv) is a consequence of two easy estimates whenever f is of moderate decrease. First, we have

$$|(f * \mathcal{P}_y)(x)| \leq C\left(\frac{1}{(1+x^2)} + \frac{y}{x^2+y^2}\right)$$

which is proved (as in the case of the heat equation) by splitting the integral $\int_{-\infty}^{\infty} f(x-t)\mathcal{P}_y(t)\,dt$ into the part where $|t| \leq |x|/2$ and the part where $|t| \geq |x|/2$. Also, we have $|(f * \mathcal{P}_y)(x)| \leq C/y$, since $\sup_x \mathcal{P}_y(x) \leq c/y$.

Using the first estimate when $|x| \geq |y|$ and the second when $|x| \leq |y|$ gives the desired decrease at infinity.

We next show that the solution is essentially unique.

Theorem 2.7 *Suppose u is continuous on the closure of the upper half-plane $\overline{\mathbb{R}^2_+}$, satisfies $\triangle u = 0$ for $(x,y) \in \mathbb{R}^2_+$, $u(x,0) = 0$, and $u(x,y)$ vanishes at infinity. Then $u = 0$.*

A simple example shows that a condition concerning the decay of u at infinity is needed: take $u(x,y) = y$. Clearly u satisfies the steady-state heat equation and vanishes on the real line, yet u is not identically zero.

The proof of the theorem relies on a basic fact about harmonic functions, which are functions satisfying $\triangle u = 0$. The fact is that the value of a harmonic function at a point equals its average value around any circle centered at that point.

Lemma 2.8 (Mean-value property) *Suppose Ω is an open set in $\mathbb{R}^2$ and let u be a function of class C^2 with $\triangle u = 0$ in Ω. If the closure of the disc centered at (x,y) and of radius R is contained in Ω, then*

$$u(x,y) = \frac{1}{2\pi}\int_0^{2\pi} u(x + r\cos\theta, y + r\sin\theta)\, d\theta$$

for all $0 \le r \le R$.

Proof. Let $U(r,\theta) = u(x + r\cos\theta, y + r\sin\theta)$. Expressing the Laplacian in polar coordinates, the equation $\triangle u = 0$ then implies

$$0 = \frac{\partial^2 U}{\partial\theta^2} + r\frac{\partial}{\partial r}\left(r\frac{\partial U}{\partial r}\right).$$

If we define $F(r) = \frac{1}{2\pi}\int_0^{2\pi} U(r,\theta)\, d\theta$, the above gives

$$r\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial r}\right) = \frac{1}{2\pi}\int_0^{2\pi} -\frac{\partial^2 U}{\partial\theta^2}(r,\theta)\, d\theta.$$

The integral of $\partial^2 U/\partial\theta^2$ over the circle vanishes since $\partial U/\partial\theta$ is periodic, hence $r\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial r}\right) = 0$, and consequently $r\partial F/\partial r$ must be constant. Evaluating this expression at $r = 0$ we find that $\partial F/\partial r = 0$. Thus F is constant, but since $F(0) = u(x,y)$, we finally find that $F(r) = u(x,y)$ for all $0 \le r \le R$, which is the mean-value property.

Finally, note that the argument above is implicit in the proof of Theorem 5.7, Chapter 2.

To prove Theorem 2.7 we argue by contradiction. Considering separately the real and imaginary parts of u, we may suppose that u itself is real-valued, and is somewhere strictly positive, say $u(x_0,y_0) > 0$ for some $x_0 \in \mathbb{R}$ and $y_0 > 0$. We shall see that this leads to a contradiction. First, since u vanishes at infinity, we can find a large semi-disc of radius R, $D_R^+ = \{(x,y) :\ x^2 + y^2 \le R,\ \ y \ge 0\}$ outside of which $u(x,y) \le \frac{1}{2}u(x_0,y_0)$. Next, since u is continuous in D_R^+, it attains its maximum M there, so there exists a point $(x_1,y_1) \in D_R^+$ with $u(x_1,y_1) = M$, while

$u(x, y) \leq M$ in the semi-disc; also, since $u(x, y) \leq \frac{1}{2}u(x_0, y_0) \leq M/2$ outside of the semi-disc, we have $u(x, y) \leq M$ throughout the entire upper half-plane. Now the mean-value property for harmonic functions implies

$$u(x_1, y_1) = \frac{1}{2\pi}\int_0^{2\pi} u(x_1 + \rho\cos\theta, y_1 + \rho\sin\theta)\, d\theta$$

whenever the circle of integration lies in the upper half-plane. In particular, this equation holds if $0 < \rho < y_1$. Since $u(x_1, y_1)$ equals the maximum value M, and $u(x_1 + \rho\cos\theta, y_1 + \rho\sin\theta) \leq M$, it follows by continuity that $u(x_1 + \rho\cos\theta, y_1 + \rho\sin\theta) = M$ on the whole circle. For otherwise $u(x, y) \leq M - \epsilon$, on an arc of length $\delta > 0$ on the circle, and this would give

$$\frac{1}{2\pi}\int_0^{2\pi} u(x_1 + \rho\cos\theta, y_1 + \rho\sin\theta)\, d\theta \leq M - \frac{\epsilon\delta}{2\pi} < M,$$

contradicting the fact that $u(x_1, y_1) = M$. Now letting $\rho \to y_1$, and using the continuity of u again, we see that this implies $u(x_1, 0) = M > 0$, which contradicts the fact that $u(x, 0) = 0$ for all x.

3 The Poisson summation formula

The definition of the Fourier transform was motivated by the desire for a continuous version of Fourier series, applicable to functions defined on the real line. We now show that there exists a further remarkable connection between the analysis of functions on the circle and related functions on $\mathbb{R}$.

Given a function $f \in \mathcal{S}(\mathbb{R})$ on the real line, we can construct a new function on the circle by the recipe

$$F_1(x) = \sum_{n=-\infty}^{\infty} f(x + n).$$

Since f is rapidly decreasing, the series converges absolutely and uniformly on every compact subset of $\mathbb{R}$, so F_1 is continuous. Note that $F_1(x + 1) = F_1(x)$ because passage from n to $n + 1$ in the above sum merely shifts the terms on the series defining $F_1(x)$. Hence F_1 is periodic with period 1. The function F_1 is called the **periodization** of f.

There is another way to arrive at a "periodic version" of f, this time by Fourier analysis. Start with the identity

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi i\xi x}\, d\xi,$$

and consider its discrete analogue, where the integral is replaced by a sum

$$F_2(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx}.$$

Once again, the sum converges absolutely and uniformly since $\hat{f}$ belongs to the Schwartz space, hence F_2 is continuous. Moreover, F_2 is also periodic of period 1 since this is the case for each one of the exponentials $e^{2\pi inx}$.

The fundamental fact is that these two approaches, which produce F_1 and F_2, actually lead to the same function.

Theorem 3.1 (Poisson summation formula) *If $f \in \mathcal{S}(\mathbb{R})$, then*

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx}.$$

In particular, setting $x = 0$ we have

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

In other words, the Fourier coefficients of the periodization of f are given precisely by the values of the Fourier transform of f on the integers.

Proof. To check the first formula it suffices, by Theorem 2.1 in Chapter 2, to show that both sides (which are continuous) have the same Fourier coefficients (viewed as functions on the circle). Clearly, the mth Fourier coefficient of the right-hand side is $\hat{f}(m)$. For the left-hand side we have

$$\begin{aligned}
\int_0^1 \left(\sum_{n=-\infty}^{\infty} f(x+n) \right) e^{-2\pi imx}\,dx &= \sum_{n=-\infty}^{\infty} \int_0^1 f(x+n)e^{-2\pi imx}\,dx \\
&= \sum_{n=-\infty}^{\infty} \int_n^{n+1} f(y)e^{-2\pi imy}\,dy \\
&= \int_{-\infty}^{\infty} f(y)e^{-2\pi imy}\,dy \\
&= \hat{f}(m),
\end{aligned}$$

where the interchange of the sum and integral is permissible since f is rapidly decreasing. This completes the proof of the theorem.

We observe that the theorem extends to the case when we merely assume that both f and $\hat{f}$ are of moderate decrease; the proof is in fact unchanged.

It turns out that the operation of periodization is important in a number of questions, even when the Poisson summation formula does not apply. We give an example by considering the elementary function $f(x) = 1/x$, $x \neq 0$. The result is that $\sum_{n=-\infty}^{\infty} 1/(x+n)$, when summed symmetrically, gives the partial fraction decomposition of the cotangent function. In fact this sum equals $\pi \cot \pi x$, when x is not an integer. Similarly with $f(x) = 1/x^2$, we get $\sum_{n=-\infty}^{\infty} 1/(x+n)^2 = \pi^2/(\sin \pi x)^2$, whenever $x \notin \mathbb{Z}$ (see Exercise 15).

3.1 Theta and zeta functions

We define the **theta function** $\vartheta(s)$ for $s > 0$ by

$$\vartheta(s) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 s}.$$

The condition on s ensures the absolute convergence of the series. A crucial fact about this special function is that it satisfies the following functional equation.

Theorem 3.2 $s^{-1/2}\vartheta(1/s) = \vartheta(s)$ *whenever* $s > 0$.

The proof of this identity consists of a simple application of the Poisson summation formula to the pair

$$f(x) = e^{-\pi s x^2} \quad \text{and} \quad \hat{f}(\xi) = s^{-1/2} e^{-\pi \xi^2/s}.$$

The theta function $\vartheta(s)$ also extends to complex values of s when $\mathrm{Re}(s) > 0$, and the functional equation is still valid then. The theta function is intimately connected with an important function in number theory, the **zeta function** $\zeta(s)$ defined for $\mathrm{Re}(s) > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Later we will see that this function carries essential information about the prime numbers (see Chapter 8).

It also turns out that ζ, ϑ, and another important function Γ are related by the following identity:

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{2}\int_0^{\infty} t^{s/2-1}(\vartheta(t) - 1)\, dt,$$

which is valid for $s > 1$ (Exercises 17 and 18).

Returning to the function ϑ, define the generalization $\Theta(z|\tau)$ given by

$$\Theta(z|\tau) = \sum_{n=-\infty}^{\infty} e^{i\pi n^2 \tau} e^{2\pi i n z}$$

whenever $\mathrm{Im}(\tau) > 0$ and $z \in \mathbb{C}$. Taking $z = 0$ and $\tau = is$ we get $\Theta(z|\tau) = \vartheta(s)$.

3.2 Heat kernels

Another application related to the Poisson summation formula and the theta function is the time-dependent heat equation on the circle. A solution to the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

subject to $u(x,0) = f(x)$, where f is periodic of period 1, was given in the previous chapter by

$$u(x,t) = (f * H_t)(x)$$

where $H_t(x)$ is the heat kernel on the circle, that is,

$$H_t(x) = \sum_{n=-\infty}^{\infty} e^{-4\pi^2 n^2 t} e^{2\pi i n x}.$$

Note in particular that with our definition of the generalized theta function in the previous section, we have $\Theta(x|4\pi i t) = H_t(x)$. Also, recall that the heat equation on $\mathbb{R}$ gave rise to the heat kernel

$$\mathcal{H}_t(x) = \frac{1}{(4\pi t)^{1/2}} e^{-x^2/4t}$$

where $\hat{\mathcal{H}}_t(\xi) = e^{-4\pi^2 \xi^2 t}$. The fundamental relation between these two objects is an immediate consequence of the Poisson summation formula:

Theorem 3.3 *The heat kernel on the circle is the periodization of the heat kernel on the real line:*

$$H_t(x) = \sum_{n=-\infty}^{\infty} \mathcal{H}_t(x+n).$$

Although the proof that $\mathcal{H}_t$ is a good kernel on $\mathbb{R}$ was fairly straightforward, we left open the harder problem that H_t is a good kernel on the circle. The above results allow us to resolve this matter.

Corollary 3.4 *The kernel $H_t(x)$ is a good kernel for $t \to 0$.*

Proof. We already observed that $\int_{|x|\leq 1/2} H_t(x)\,dx = 1$. Now note that $H_t \geq 0$, which is immediate from the above formula since $\mathcal{H}_t \geq 0$. Finally, we claim that when $|x| \leq 1/2$,

$$H_t(x) = \mathcal{H}_t(x) + \mathcal{E}_t(x),$$

where the error satisfies $|\mathcal{E}_t(x)| \leq c_1 e^{-c_2/t}$ with $c_1, c_2 > 0$ and $0 < t \leq 1$. To see this, note again that the formula in the theorem gives

$$H_t(x) = \mathcal{H}_t(x) + \sum_{|n|\geq 1} \mathcal{H}_t(x+n);$$

therefore, since $|x| \leq 1/2$,

$$\mathcal{E}_t(x) = \frac{1}{\sqrt{4\pi t}} \sum_{|n|\geq 1} e^{-(x+n)^2/4t} \leq Ct^{-1/2} \sum_{n\geq 1} e^{-cn^2/t}.$$

Note that $n^2/t \geq n^2$ and $n^2/t \geq 1/t$ whenever $0 < t \leq 1$, so $e^{-cn^2/t} \leq e^{-\frac{c}{2}n^2} e^{-\frac{c}{2}\frac{1}{t}}$. Hence

$$|\mathcal{E}_t(x)| \leq Ct^{-1/2} e^{-\frac{c}{2}\frac{1}{t}} \sum_{n\geq 1} e^{-\frac{c}{2}n^2} \leq c_1 e^{-c_2/t}.$$

The proof of the claim is complete, and as a result $\int_{|x|\leq 1/2} |\mathcal{E}_t(x)|\,dx \to 0$ as $t \to 0$. It is now clear that H_t satisfies

$$\int_{\eta<|x|\leq 1/2} |H_t(x)|\,dx \to 0 \quad \text{as } t \to 0,$$

because $\mathcal{H}_t$ does.

3.3 Poisson kernels

In a similar manner to the discussion above about the heat kernels, we state the relation between the Poisson kernels for the disc and the upper half-plane where

$$P_r(\theta) = \frac{1-r^2}{1-2r\cos\theta + r^2} \quad \text{and} \quad \mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{y^2+x^2}.$$

Theorem 3.5 $P_r(2\pi x) = \sum_{n\in\mathbb{Z}} \mathcal{P}_y(x+n)$ *where* $r = e^{-2\pi y}$.

This is again an immediate corollary of the Poisson summation formula applied to $f(x) = \mathcal{P}_y(x)$ and $\hat{f}(\xi) = e^{-2\pi|\xi|y}$. Of course, here we use the Poisson summation formula under the assumptions that f and $\hat{f}$ are of moderate decrease.

4 The Heisenberg uncertainty principle

The mathematical thrust of the principle can be formulated in terms of a relation between a function and its Fourier transform. The basic underlying law, formulated in its vaguest and most general form, states that a function and its Fourier transform cannot both be essentially localized. Somewhat more precisely, if the "preponderance" of the mass of a function is concentrated in an interval of length L, then the preponderance of the mass of its Fourier transform cannot lie in an interval of length essentially smaller than L^{-1}. The exact statement is as follows.

Theorem 4.1 *Suppose ψ is a function in $\mathcal{S}(\mathbb{R})$ which satisfies the normalizing condition $\int_{-\infty}^{\infty} |\psi(x)|^2\,dx = 1$. Then*

$$\left(\int_{-\infty}^{\infty} x^2|\psi(x)|^2\,dx\right)\left(\int_{-\infty}^{\infty} \xi^2|\hat{\psi}(\xi)|^2\,d\xi\right) \geq \frac{1}{16\pi^2},$$

and equality holds if and only if $\psi(x) = Ae^{-Bx^2}$ where $B > 0$ and $|A|^2 = \sqrt{2B/\pi}$.

In fact, we have

$$\left(\int_{-\infty}^{\infty} (x-x_0)^2|\psi(x)|^2\,dx\right)\left(\int_{-\infty}^{\infty} (\xi-\xi_0)^2|\hat{\psi}(\xi)|^2\,d\xi\right) \geq \frac{1}{16\pi^2}$$

for every $x_0, \xi_0 \in \mathbb{R}$.

Proof. The second inequality actually follows from the first by replacing $\psi(x)$ by $e^{-2\pi i x\xi_0}\psi(x+x_0)$ and changing variables. To prove the first inequality, we argue as follows. Beginning with our normalizing assumption $\int |\psi|^2 = 1$, and recalling that ψ and ψ' are rapidly decreasing, an integration by parts gives

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} |\psi(x)|^2\,dx \\
&= -\int_{-\infty}^{\infty} x\frac{d}{dx}|\psi(x)|^2\,dx \\
&= -\int_{-\infty}^{\infty} \left(x\psi'(x)\overline{\psi(x)} + x\overline{\psi'(x)}\psi(x)\right)\,dx.
\end{aligned}$$

The last identity follows because $|\psi|^2 = \psi\overline{\psi}$. Therefore

$$\begin{aligned} 1 &\leq 2\int_{-\infty}^{\infty} |x|\,|\psi(x)|\,|\psi'(x)|\,dx \\ &\leq 2\left(\int_{-\infty}^{\infty} x^2|\psi(x)|^2\,dx\right)^{1/2}\left(\int_{-\infty}^{\infty} |\psi'(x)|^2\,dx\right)^{1/2}, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality. The identity

$$\int_{-\infty}^{\infty} |\psi'(x)|^2\,dx = 4\pi^2\int_{-\infty}^{\infty} \xi^2|\hat{\psi}(\xi)|^2\,d\xi,$$

which holds because of the properties of the Fourier transform and the Plancherel formula, concludes the proof of the inequality in the theorem.

If equality holds, then we must also have equality where we applied the Cauchy-Schwarz inequality, and as a result we find that $\psi'(x) = \beta x\psi(x)$ for some constant β. The solutions to this equation are $\psi(x) = Ae^{\beta x^2/2}$, where A is constant. Since we want ψ to be a Schwartz function, we must take $\beta = -2B < 0$, and since we impose the condition $\int_{-\infty}^{\infty} |\psi(x)|^2\,dx = 1$ we find that $|A|^2 = \sqrt{2B/\pi}$, as was to be shown.

The precise assertion contained in Theorem 4.1 first came to light in the study of quantum mechanics. It arose when one considered the extent to which one could simultaneously locate the position and momentum of a particle. Assuming we are dealing with (say) an electron that travels along the real line, then according to the laws of physics, matters are governed by a "state function" ψ, which we can assume to be in $\mathcal{S}(\mathbb{R})$, and which is normalized according to the requirement that

$$\int_{-\infty}^{\infty} |\psi(x)|^2\,dx = 1. \tag{11}$$

The position of the particle is then determined not as a definite point x; instead its probable location is given by the rules of quantum mechanics as follows:

- *The probability that the particle is located in the interval* (a, b) *is* $\int_a^b |\psi(x)|^2\,dx$.

According to this law we can calculate the probable location of the particle with the aid of ψ: in fact, there may be only a small probability that the particle is located in a given interval (a', b'), but nevertheless it is somewhere on the real line since $\int_{-\infty}^{\infty} |\psi(x)|^2\,dx = 1$.

In addition to the **probability density** $|\psi(x)|^2 dx$, there is the **expectation** of where the particle might be. This expectation is the best guess of the position of the particle, given its probability distribution determined by $|\psi(x)|^2 dx$, and is the quantity defined by

$$\overline{x} = \int_{-\infty}^{\infty} x|\psi(x)|^2 \, dx. \tag{12}$$

Why is this our best guess? Consider the simpler (idealized) situation where we are given that the particle can be found at only finitely many different points, $x_1, x_2, \ldots, x_N$ on the real axis, with p_i the probability that the particle is at x_i, and $p_1 + p_2 + \cdots + p_N = 1$. Then, if we knew nothing else, and were forced to make one choice as to the position of the particle, we would naturally take $\overline{x} = \sum_{i=1}^{N} x_i p_i$, which is the appropriate weighted average of the possible positions. The quantity (12) is clearly the general (integral) version of this.

We next come to the notion of **variance**, which in our terminology is the **uncertainty** attached to our expectation. Having determined that the expected position of the particle is $\overline{x}$ (given by (12)), the resulting uncertainty is the quantity

$$\int_{-\infty}^{\infty} (x - \overline{x})^2 |\psi(x)|^2 \, dx. \tag{13}$$

Notice that if ψ is highly concentrated near $\overline{x}$, it means that there is a high probability that x is near $\overline{x}$, and so (13) is small, because most of the contribution to the integral takes place for values of x near $\overline{x}$. Here we have a small uncertainty. On the other hand, if $\psi(x)$ is rather flat (that is, the probability distribution $|\psi(x)|^2 dx$ is not very concentrated), then the integral (13) is rather big, because large values of $(x - \overline{x})^2$ will come into play, and as a result the uncertainty is relatively large.

It is also worthwhile to observe that the expectation $\overline{x}$ is that choice for which the uncertainty $\int_{-\infty}^{\infty} (x - \overline{x})^2 |\psi(x)|^2 \, dx$ is the smallest. Indeed, if we try to minimize this quantity by equating to 0 its derivative with respect to $\overline{x}$, we find that $2\int_{-\infty}^{\infty} (x - \overline{x})|\psi(x)|^2 \, dx = 0$, which gives (12).

So far, we have discussed the "expectation" and "uncertainty" related to the position of the particle. Of equal relevance are the corresponding notions regarding its momentum. The corresponding rule of quantum mechanics is:

- *The probability that the momentum ξ of the particle belongs to the interval (a, b) is $\int_a^b |\hat{\psi}(\xi)|^2 \, d\xi$ where $\hat{\psi}$ is the Fourier transform of ψ.*

Combining these two laws with Theorem 4.1 gives $1/16\pi^2$ as the lower bound for the product of the uncertainty of the position and the uncertainty of the momentum of a particle. So the more certain we are about the location of the particle, the less certain we can be about its momentum, and vice versa. However, we have simplified the statement of the two laws by rescaling to change the units of measurement. Actually, there enters a fundamental but small physical number $\hbar$ called Planck's constant. When properly taken into account, the physical conclusion is

$$\text{(uncertainty of position)}\times\text{(uncertainty of momentum)} \geq \hbar/16\pi^2.$$

5 Exercises

1. Corollary 2.3 in Chapter 2 leads to the following simplified version of the Fourier inversion formula. Suppose f is a continuous function supported on an interval $[-M, M]$, whose Fourier transform $\hat{f}$ is of moderate decrease.

(a) Fix L with $L/2 > M$, and show that $f(x) = \sum a_n(L)e^{2\pi inx/L}$ where

$$a_n(L) = \frac{1}{L}\int_{-L/2}^{L/2} f(x)e^{-2\pi inx/L}\,dx = \frac{1}{L}\hat{f}(n/L).$$

Alternatively, we may write $f(x) = \delta\sum_{n=-\infty}^{\infty}\hat{f}(n\delta)e^{2\pi in\delta x}$ with $\delta = 1/L$.

(b) Prove that if F is continuous and of moderate decrease, then

$$\int_{-\infty}^{\infty} F(\xi)\,d\xi = \lim_{\substack{\delta\to 0\\ \delta>0}} \delta\sum_{n=-\infty}^{\infty} F(\delta n).$$

(c) Conclude that $f(x) = \displaystyle\int_{-\infty}^{\infty}\hat{f}(\xi)e^{2\pi ix\xi}\,d\xi$.

[Hint: For (a), note that the Fourier series of f on $[-L/2, L/2]$ converges absolutely. For (b), first approximate the integral by $\int_{-N}^{N} F$ and the sum by $\delta\sum_{|n|\leq N/\delta} F(n\delta)$. Then approximate the second integral by Riemann sums.]

2. Let f and g be the functions defined by

$$f(x) = \chi_{[-1,1]}(x) = \begin{cases} 1 & \text{if } |x| \leq 1, \\ 0 & \text{otherwise,}\end{cases} \quad \text{and} \quad g(x) = \begin{cases} 1-|x| & \text{if } |x|\leq 1, \\ 0 & \text{otherwise.}\end{cases}$$

Although f is not continuous, the integral defining its Fourier transform still makes sense. Show that

$$\hat{f}(\xi) = \frac{\sin 2\pi\xi}{\pi\xi} \quad \text{and} \quad \hat{g}(\xi) = \left(\frac{\sin\pi\xi}{\pi\xi}\right)^2,$$

with the understanding that $\hat{f}(0) = 2$ and $\hat{g}(0) = 1$.

3. The following exercise illustrates the principle that the decay of $\hat{f}$ is related to the continuity properties of f.

(a) Suppose that f is a function of moderate decrease on $\mathbb{R}$ whose Fourier transform $\hat{f}$ is continuous and satisfies

$$\hat{f}(\xi) = O\left(\frac{1}{|\xi|^{1+\alpha}}\right) \qquad \text{as } |\xi| \to \infty$$

for some $0 < \alpha < 1$. Prove that f satisfies a Hölder condition of order α, that is, that

$$|f(x+h) - f(x)| \leq M|h|^{\alpha} \qquad \text{for some } M > 0 \text{ and all } x, h \in \mathbb{R}.$$

(b) Let f be a continuous function on $\mathbb{R}$ which vanishes for $|x| \geq 1$, with $f(0) = 0$, and which is equal to $1/\log(1/|x|)$ for all x in a neighborhood of the origin. Prove that $\hat{f}$ is not of moderate decrease. In fact, there is no $\epsilon > 0$ so that $\hat{f}(\xi) = O(1/|\xi|^{1+\epsilon})$ as $|\xi| \to \infty$.

[Hint: For part (a), use the Fourier inversion formula to express $f(x+h) - f(x)$ as an integral involving $\hat{f}$, and estimate this integral separately for ξ in the two ranges $|\xi| \leq 1/|h|$ and $|\xi| \geq 1/|h|$.]

4. Bump functions. Examples of compactly supported functions in $\mathcal{S}(\mathbb{R})$ are very handy in many applications in analysis. Some examples are:

(a) Suppose $a < b$, and f is the function such that $f(x) = 0$ if $x \leq a$ or $x \geq b$ and

$$f(x) = e^{-1/(x-a)}e^{-1/(b-x)} \qquad \text{if } a < x < b.$$

Show that f is indefinitely differentiable on $\mathbb{R}$.

(b) Prove that there exists an indefinitely differentiable function F on $\mathbb{R}$ such that $F(x) = 0$ if $x \leq a$, $F(x) = 1$ if $x \geq b$, and F is strictly increasing on $[a, b]$.

(c) Let $\delta > 0$ be so small that $a + \delta < b - \delta$. Show that there exists an indefinitely differentiable function g such that g is 0 if $x \leq a$ or $x \geq b$, g is 1 on $[a + \delta, b - \delta]$, and g is strictly monotonic on $[a, a + \delta]$ and $[b - \delta, b]$.

[Hint: For (b) consider $F(x) = c\int_{-\infty}^{x} f(t)\,dt$ where c is an appropriate constant.]

5. Suppose f is continuous and of moderate decrease.

(a) Prove that $\hat{f}$ is continuous and $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$.

(b) Show that if $\hat{f}(\xi) = 0$ for all ξ, then f is identically 0.

[Hint: For part (a), show that $\hat{f}(\xi) = \frac{1}{2}\int_{-\infty}^{\infty}[f(x) - f(x - 1/(2\xi))]e^{-2\pi i x\xi}\,dx$. For part (b), verify that the multiplication formula $\int f(x)\hat{g}(x)\,dx = \int \hat{f}(y)g(y)\,dy$ still holds whenever $g \in \mathcal{S}(\mathbb{R})$.]

6. The function $e^{-\pi x^2}$ is its own Fourier transform. Generate other functions that (up to a constant multiple) are their own Fourier transforms. What must the constant multiples be? To decide this, prove that $\mathcal{F}^4 = I$. Here $\mathcal{F}(f) = \hat{f}$ is the Fourier transform, $\mathcal{F}^4 = \mathcal{F} \circ \mathcal{F} \circ \mathcal{F} \circ \mathcal{F}$, and I is the identity operator $(If)(x) = f(x)$ (see also Problem 7).

7. Prove that the convolution of two functions of moderate decrease is a function of moderate decrease.

[Hint: Write

$$\int f(x-y)g(y)\,dy = \int_{|y|\leq |x|/2} + \int_{|y|\geq |x|/2}.$$

In the first integral $f(x-y) = O(1/(1+x^2))$ while in the second integral $g(y) = O(1/(1+x^2))$.]

8. Prove that f is continuous, of moderate decrease, and $\int_{-\infty}^{\infty} f(y)e^{-y^2}e^{2xy}dy = 0$ for all $x \in \mathbb{R}$, then $f = 0$.

[Hint: Consider $f * e^{-x^2}$.]

9. If f is of moderate decrease, then

$$\int_{-R}^{R}\left(1 - \frac{|\xi|}{R}\right)\hat{f}(\xi)e^{2\pi i x\xi}\,d\xi = (f * \mathcal{F}_R)(x), \tag{14}$$

where the Fejér kernel on the real line is defined by

$$\mathcal{F}_R(t) = \begin{cases} R\left(\dfrac{\sin \pi t R}{\pi t R}\right)^2 & \text{if } t \neq 0, \\ R & \text{if } t = 0. \end{cases}$$

Show that $\{\mathcal{F}_R\}$ is a family of good kernels as $R \to \infty$, and therefore (14) tends uniformly to $f(x)$ as $R \to \infty$. This is the analogue of Fejér's theorem for Fourier series in the context of the Fourier transform.

10. Below is an outline of a different proof of the Weierstrass approximation theorem.

Define the **Landau** kernels by

$$L_n(x) = \begin{cases} \dfrac{(1-x^2)^n}{c_n} & \text{if } -1 \le x \le 1, \\ 0 & \text{if } |x| \ge 1, \end{cases}$$

where c_n is chosen so that $\int_{-\infty}^{\infty} L_n(x)\,dx = 1$. Prove that $\{L_n\}_{n\ge 0}$ is a family of good kernels as $n \to \infty$. As a result, show that if f is a continuous function supported in $[-1/2, 1/2]$, then $(f * L_n)(x)$ is a sequence of polynomials on $[-1/2, 1/2]$ which converges uniformly to f.
[Hint: First show that $c_n \ge 2/(n+1)$.]

11. Suppose that u is the solution to the heat equation given by $u = f * \mathcal{H}_t$ where $f \in \mathcal{S}(\mathbb{R})$. If we also set $u(x, 0) = f(x)$, prove that u is continuous on the closure of the upper half-plane, and vanishes at infinity, that is,

$$u(x,t) \to 0 \qquad \text{as } |x| + t \to \infty.$$

[Hint: To prove that u vanishes at infinity, show that (i) $|u(x,t)| \le C/\sqrt{t}$ and (ii) $|u(x,t)| \le C/(1+|x|^2) + Ct^{-1/2}e^{-cx^2/t}$. Use (i) when $|x| \le t$, and (ii) otherwise.]

12. Show that the function defined by

$$u(x,t) = \frac{x}{t}\mathcal{H}_t(x)$$

satisfies the heat equation for $t > 0$ and $\lim_{t\to 0} u(x,t) = 0$ for every x, but u is *not* continuous at the origin.
[Hint: Approach the origin with (x,t) on the parabola $x^2/4t = c$ where c is a constant.]

13. Prove the following uniqueness theorem for harmonic functions in the strip $\{(x,y) : 0 < y < 1,\ -\infty < x < \infty\}$: if u is harmonic in the strip, continuous on its closure with $u(x,0) = u(x,1) = 0$ for all $x \in \mathbb{R}$, and u vanishes at infinity, then $u = 0$.

14. Prove that the periodization of the Fejér kernel $\mathcal{F}_N$ on the real line (Exercise 9) is equal to the Fejér kernel for periodic functions of period 1. In other words,

$$\sum_{n=-\infty}^{\infty} \mathcal{F}_N(x+n) = F_N(x),$$

when $N \ge 1$ is an integer, and where

$$F_N(x) = \sum_{n=-N}^{N}\left(1 - \frac{|n|}{N}\right)e^{2\pi i n x} = \frac{1}{N}\frac{\sin^2(N\pi x)}{\sin^2(\pi x)}.$$

15. This exercise provides another example of periodization.

(a) Apply the Poisson summation formula to the function g in Exercise 2 to obtain

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\alpha)^2} = \frac{\pi^2}{(\sin \pi\alpha)^2}$$

whenever α is real, but not equal to an integer.

(b) Prove as a consequence that

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\alpha)} = \frac{\pi}{\tan \pi\alpha} \tag{15}$$

whenever α is real but not equal to an integer. [Hint: First prove it when $0 < \alpha < 1$. To do so, integrate the formula in (b). What is the precise meaning of the series on the left-hand side of (15)? Evaluate at $\alpha = 1/2$.]

16. The Dirichlet kernel on the real line is defined by

$$\int_{-R}^{R} \hat{f}(\xi) e^{2\pi i x\xi}\, d\xi = (f * \mathcal{D}_R)(x) \quad \text{so that} \quad \mathcal{D}_R(x) = \widehat{\chi_{[-R,R]}}(x) = \frac{\sin(2\pi R x)}{\pi x}.$$

Also, the modified Dirichlet kernel for periodic functions of period 1 is defined by

$$D_N^*(x) = \sum_{|n|\leq N-1} e^{2\pi i n x} + \frac{1}{2}(e^{-2\pi i N x} + e^{2\pi i N x}).$$

Show that the result in Exercise 15 gives

$$\sum_{n=-\infty}^{\infty} \mathcal{D}_N(x+n) = D_N^*(x),$$

where $N \geq 1$ is an integer, and the infinite series must be summed symmetrically. In other words, the periodization of $\mathcal{D}_N$ is the modified Dirichlet kernel D_N^*.

17. The **gamma function** is defined for $s > 0$ by

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1}\, dx.$$

(a) Show that for $s > 0$ the above integral makes sense, that is, that the following two limits exist:

$$\lim_{\substack{\delta \to 0 \\ \delta > 0}} \int_\delta^1 e^{-x} x^{s-1}\, dx \quad \text{and} \quad \lim_{A\to\infty} \int_1^A e^{-x} x^{s-1}\, dx.$$

(b) Prove that $\Gamma(s+1) = s\Gamma(s)$ whenever $s > 0$, and conclude that for every integer $n \geq 1$ we have $\Gamma(n+1) = n!$.

(c) Show that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad \text{and} \quad \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

[Hint: For (c), use $\int_{-\infty}^{\infty} e^{-\pi x^2}\, dx = 1$.]

18. The **zeta function** is defined for $s > 1$ by $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$. Verify the identity

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{2}\int_0^{\infty} t^{\frac{s}{2}-1}(\vartheta(t) - 1)\, dt \qquad \text{whenever } s > 1$$

where Γ and ϑ are the gamma and theta functions, respectively:

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1}\, dt \quad \text{and} \quad \vartheta(s) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 s}.$$

More about the zeta function and its relation to the prime number theorem can be found in Book II.

19. The following is a variant of the calculation of $\zeta(2m) = \sum_{n=1}^{\infty} 1/n^{2m}$ found in Problem 4, Chapter 3.

(a) Apply the Poisson summation formula to $f(x) = t/(\pi(x^2 + t^2))$ and $\hat{f}(\xi) = e^{-2\pi t|\xi|}$ where $t > 0$ in order to get

$$\frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{t}{t^2 + n^2} = \sum_{n=-\infty}^{\infty} e^{-2\pi t|n|}.$$

(b) Prove the following identity valid for $0 < t < 1$:

$$\frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{t}{t^2 + n^2} = \frac{1}{\pi t} + \frac{2}{\pi}\sum_{m=1}^{\infty}(-1)^{m+1}\zeta(2m)t^{2m-1}$$

as well as

$$\sum_{n=-\infty}^{\infty} e^{-2\pi t|n|} = \frac{2}{1 - e^{-2\pi t}} - 1.$$

(c) Use the fact that

$$\frac{z}{e^z - 1} = 1 - \frac{z}{2} + \sum_{m=1}^{\infty} \frac{B_{2m}}{(2m)!} z^{2m},$$

where B_k are the Bernoulli numbers to deduce from the above formula,

$$2\zeta(2m) = (-1)^{m+1} \frac{(2\pi)^{2m}}{(2m)!} B_{2m}.$$

20. The following results are relevant in information theory when one tries to recover a signal from its samples.

Suppose f is of moderate decrease and that its Fourier transform $\hat{f}$ is supported in $I = [-1/2, 1/2]$. Then, f is entirely determined by its restriction to $\mathbb{Z}$. This means that if g is another function of moderate decrease whose Fourier transform is supported in I and $f(n) = g(n)$ for all $n \in \mathbb{Z}$, then $f = g$. More precisely:

(a) Prove that the following reconstruction formula holds:

$$f(x) = \sum_{n=-\infty}^{\infty} f(n) K(x - n) \qquad \text{where } K(y) = \frac{\sin \pi y}{\pi y}.$$

Note that $K(y) = O(1/|y|)$ as $|y| \to \infty$.

(b) If $\lambda > 1$, then

$$f(x) = \sum_{n=-\infty}^{\infty} \frac{1}{\lambda} f\left(\frac{n}{\lambda}\right) K_\lambda\left(x - \frac{n}{\lambda}\right) \qquad \text{where } K_\lambda(y) = \frac{\cos \pi y - \cos \pi \lambda y}{\pi^2 (\lambda - 1) y^2}.$$

Thus, if one samples f "more often," the series in the reconstruction formula converges faster since $K_\lambda(y) = O(1/|y|^2)$ as $|y| \to \infty$. Note that $K_\lambda(y) \to K(y)$ as $\lambda \to 1$.

(c) Prove that $\displaystyle\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |f(n)|^2.$

[Hint: For part (a) show that if χ is the characteristic function of I, then $\hat{f}(\xi) = \chi(\xi) \sum_{n=-\infty}^{\infty} f(n) e^{-2\pi i n \xi}$. For (b) use the function in Figure 2 instead of $\chi(\xi)$.]

21. Suppose that f is continuous on $\mathbb{R}$. Show that f and $\hat{f}$ cannot both be compactly supported unless $f = 0$. This can be viewed in the same spirit as the uncertainty principle.

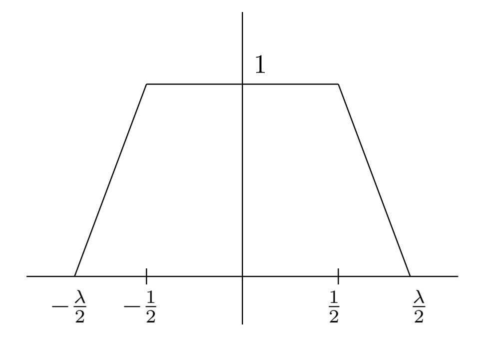


Figure 2. The function in Exercise 20

[Hint: Assume f is supported in $[0, 1/2]$. Expand f in a Fourier series in the interval $[0, 1]$, and note that as a result, f is a trigonometric polynomial.]

22. The heuristic assertion stated before Theorem 4.1 can be made precise as follows. If F is a function on $\mathbb{R}$, then we say that the preponderance of its mass is contained in an interval I (centered at the origin) if

$$\int_I x^2|F(x)|^2\,dx \geq \frac{1}{2}\int_{\mathbb{R}} x^2|F(x)|^2\,dx. \tag{16}$$

Now suppose $f \in \mathcal{S}$, and (16) holds with $F = f$ and $I = I_1$; also with $F = \hat{f}$ and $I = I_2$. Then if L_j denotes the length of I_j, we have

$$L_1 L_2 \geq \frac{1}{2\pi}.$$

A similar conclusion holds if the intervals are not necessarily centered at the origin.

23. The Heisenberg uncertainty principle can be formulated in terms of the operator $L = -\frac{d^2}{dx^2} + x^2$, which acts on Schwartz functions by the formula

$$L(f) = -\frac{d^2 f}{dx^2} + x^2 f.$$

This operator, sometimes called the **Hermite operator**, is the quantum analogue of the harmonic oscillator. Consider the usual inner product on $\mathcal{S}$ given by

$$(f, g) = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\,dx \qquad \text{whenever } f, g \in \mathcal{S}.$$

(a) Prove that the Heisenberg uncertainty principle implies

$$(Lf, f) \geq (f, f) \qquad \text{for all } f \in \mathcal{S}.$$

This is usually denoted by $L \geq I$. [Hint: Integrate by parts.]

(b) Consider the operators A and A^* defined on $\mathcal{S}$ by

$$A(f) = \frac{df}{dx} + xf \qquad \text{and} \qquad A^*(f) = -\frac{df}{dx} + xf.$$

The operators A and A^* are sometimes called the **annihilation** and **creation** operators, respectively. Prove that for all $f, g \in \mathcal{S}$ we have

(i) $(Af, g) = (f, A^*g)$,

(ii) $(Af, Af) = (A^*Af, f) \geq 0$,

(iii) $A^*A = L - I$.

In particular, this again shows that $L \geq I$.

(c) Now for $t \in \mathbb{R}$, let

$$A_t(f) = \frac{df}{dx} + txf \qquad \text{and} \qquad A_t^*(f) = -\frac{df}{dx} + txf.$$

Use the fact that $(A_t^* A_t f, f) \geq 0$ to give another proof of the Heisenberg uncertainty principle which says that whenever $\int_{-\infty}^{\infty} |f(x)|^2\, dx = 1$ then

$$\left(\int_{-\infty}^{\infty} x^2 |f(x)|^2\, dx \right) \left(\int_{-\infty}^{\infty} \left| \frac{df}{dx} \right|^2 dx \right) \geq 1/4.$$

[Hint: Think of $(A_t^* A_t f, f)$ as a quadratic polynomial in t.]

6 Problems

1. The equation

$$x^2 \frac{\partial^2 u}{\partial x^2} + ax \frac{\partial u}{\partial x} = \frac{\partial u}{\partial t} \tag{17}$$

with $u(x, 0) = f(x)$ for $0 < x < \infty$ and $t > 0$ is a variant of the heat equation which occurs in a number of applications. To solve (17), make the change of variables $x = e^{-y}$ so that $-\infty < y < \infty$. Set $U(y, t) = u(e^{-y}, t)$ and $F(y) = f(e^{-y})$. Then the problem reduces to the equation

$$\frac{\partial^2 U}{\partial y^2} + (1 - a) \frac{\partial U}{\partial y} = \frac{\partial U}{\partial t},$$

with $U(y,0) = F(y)$. This can be solved like the usual heat equation (the case $a = 1$) by taking the Fourier transform in the y variable. One must then compute the integral $\int_{-\infty}^{\infty} e^{(-4\pi^2\xi^2+(1-a)2\pi i\xi)t} e^{2\pi i\xi v}\, d\xi$. Show that the solution of the original problem is then given by

$$u(x,t) = \frac{1}{(4\pi t)^{1/2}} \int_0^{\infty} e^{-(\log(v/x)+(1-a)t)^2/(4t)} f(v)\, \frac{dv}{v}.$$

2. The **Black-Scholes** equation from finance theory is

$$\frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} - rV = 0, \qquad 0 < t < T, \tag{18}$$

subject to the "final" boundary condition $V(s,T) = F(s)$. An appropriate change of variables reduces this to the equation in Problem 1. Alternatively, the substitution $V(s,t) = e^{ax+b\tau}U(x,\tau)$ where $x = \log s$, $\tau = \frac{\sigma^2}{2}(T-t)$, $a = \frac{1}{2} - \frac{r}{\sigma^2}$, and $b = -\left(\frac{1}{2} + \frac{r}{\sigma^2}\right)^2$ reduces (18) to the one-dimensional heat equation with the initial condition $U(x,0) = e^{-ax}F(e^x)$. Thus a solution to the Black-Scholes equation is

$$V(s,t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^{\infty} e^{-\frac{(\log(s/s^*)+(r-\sigma^2/2)(T-t))^2}{2\sigma^2(T-t)}} F(s^*)\, ds^*.$$

3. * **The Dirichlet problem in a strip**. Consider the equation $\triangle u = 0$ in the horizontal strip

$$\{(x,y) : 0 < y < 1,\ -\infty < x < \infty\}$$

with boundary conditions $u(x,0) = f_0(x)$ and $u(x,1) = f_1(x)$, where f_0 and f_1 are both in the Schwartz space.

(a) Show (formally) that if u is a solution to this problem, then

$$\hat{u}(\xi,y) = A(\xi)e^{2\pi\xi y} + B(\xi)e^{-2\pi\xi y}.$$

Express A and B in terms of $\widehat{f_0}$ and $\widehat{f_1}$, and show that

$$\hat{u}(\xi,y) = \frac{\sinh(2\pi(1-y)\xi)}{\sinh(2\pi\xi)}\widehat{f_0}(\xi) + \frac{\sinh(2\pi y\xi)}{\sinh(2\pi\xi)}\widehat{f_0}(\xi).$$

(b) Prove as a result that

$$\int_{-\infty}^{\infty} |u(x,y) - f_0(x)|^2\, dx \to 0 \qquad \text{as } y \to 0$$

and

$$\int_{-\infty}^{\infty} |u(x,y) - f_1(x)|^2\, dx \to 0 \qquad \text{as } y \to 1.$$

(c) If $\Phi(\xi) = (\sinh 2\pi a\xi)/(\sinh 2\pi\xi)$, with $0 \leq a < 1$, then Φ is the Fourier transform of φ where

$$\varphi(x) = \frac{\sin \pi a}{2} \cdot \frac{1}{\cosh \pi x + \cos \pi a}.$$

This can be shown, for instance, by using contour integration and the residue formula from complex analysis (see Book II, Chapter 3).

(d) Use this result to express u in terms of Poisson-like integrals involving f_0 and f_1 as follows:

$$u(x, y) = \frac{\sin \pi y}{2} \left(\int_{-\infty}^{\infty} \frac{f_0(x-t)}{\cosh \pi t - \cos \pi y}\, dt + \int_{-\infty}^{\infty} \frac{f_1(x-t)}{\cosh \pi t + \cos \pi y}\, dt \right).$$

(e) Finally, one can check that the function $u(x, y)$ defined by the above expression is harmonic in the strip, and converges uniformly to $f_0(x)$ as $y \to 0$, and to $f_1(x)$ as $y \to 1$. Moreover, one sees that $u(x, y)$ vanishes at infinity, that is, $\lim_{|x|\to\infty} u(x, y) = 0$, uniformly in y.

In Exercise 12, we gave an example of a function that satisfies the heat equation in the upper half-plane, with boundary value 0, but which was not identically 0. We observed in this case that u was in fact not continuous up to the boundary.

In Problem 4 we exhibit examples illustrating non-uniqueness, but this time with continuity up to the boundary $t = 0$. These examples satisfy a growth condition at infinity, namely $|u(x,t)| \leq Ce^{cx^{2+\epsilon}}$, for any $\epsilon > 0$. Problems 5 and 6 show that under the more restrictive growth condition $|u(x,t)| \leq Ce^{cx^2}$, uniqueness does hold.

4.* If g is a smooth function on $\mathbb{R}$, define the formal power series

$$u(x, t) = \sum_{n=0}^{\infty} g^{(n)}(t) \frac{x^{2n}}{(2n!)}. \tag{19}$$

(a) Check formally that u solves the heat equation.

(b) For $a > 0$, consider the function defined by

$$g(t) = \begin{cases} e^{-t^{-a}} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0. \end{cases}$$

One can show that there exists $0 < \theta < 1$ depending on a so that

$$|g^{(k)}(t)| \leq \frac{k!}{(\theta t)^k} e^{-\frac{1}{2}t^{-a}} \quad \text{for } t > 0.$$

(c) As a result, for each x and t the series (19) converges; u solves the heat equation; u vanishes for $t = 0$; and u satisfies the estimate $|u(x,t)| \leq Ce^{c|x|^{2a/(a-1)}}$ for some constants $C, c > 0$.

(d) Conclude that for every $\epsilon > 0$ there exists a non-zero solution to the heat equation which is continuous for $x \in \mathbb{R}$ and $t \geq 0$, which satisfies $u(x, 0) = 0$ and $|u(x, t)| \leq Ce^{c|x|^{2+\epsilon}}$.

5.* The following "maximum principle" for solutions of the heat equation will be used in the next problem.

Theorem. *Suppose that $u(x,t)$ is a real-valued solution of the heat equation in the upper half-plane, which is continuous on its closure. Let R denote the rectangle*

$$R = \{(x, y) \in \mathbb{R}^2 : \; a \leq x \leq b, \; 0 \leq t \leq c\}$$

and $\partial' R$ be the part of the boundary of R which consists of the two vertical sides and its base on the line $t = 0$ (see Figure 3*). Then*

$$\min_{(x,t)\in\partial' R} u(x,t) = \min_{(x,t)\in R} u(x,t) \quad \text{and} \quad \max_{(x,t)\in\partial' R} u(x,t) = \max_{(x,t)\in R} u(x,t).$$

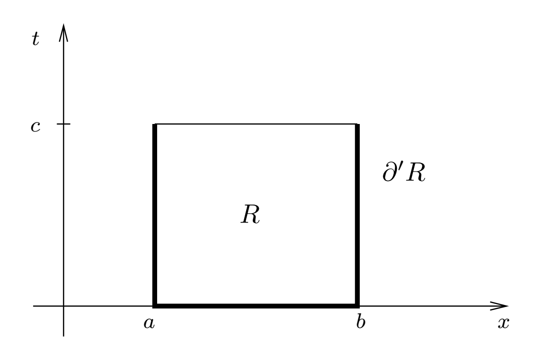


Figure 3. The rectangle R and part of its boundary $\partial' R$

The steps leading to a proof of this result are outlined below.

(a) Show that it suffices to prove that if $u \geq 0$ on $\partial' R$, then $u \geq 0$ in R.

(b) For $\epsilon > 0$, let

$$v(x, t) = u(x, t) + \epsilon t.$$

Then, v has a minimum on R, say at (x_1, t_1). Show that $x_1 = a$ or b, or else $t_1 = 0$. To do so, suppose on the contrary that $a < x_1 < b$ and $0 < t_1 \leq c$, and prove that $v_{xx}(x_1, t_1) - v_t(x_1, t_1) \leq -\epsilon$. However, show also that the left-hand side must be non-negative.

(c) Deduce from (b) that $u(x, t) \geq \epsilon(t_1 - t)$ for any $(x, t) \in R$ and let $\epsilon \to 0$.

6.* The examples in Problem 4 are optimal in the sense of the following uniqueness theorem due to Tychonoff.

Theorem. *Suppose $u(x,t)$ satisfies the following conditions:*

(i) $u(x,t)$ *solves the heat equation for all* $x \in \mathbb{R}$ *and and all* $t > 0$.

(ii) $u(x,t)$ *is continuous for all* $x \in \mathbb{R}$ *and* $0 \leq t \leq c$.

(iii) $u(x,0) = 0$.

(iv) $|u(x,t)| \leq Me^{ax^2}$ *for some* M, a, *and all* $x \in \mathbb{R}$, $0 \leq t < c$.

Then u *is identically equal to* 0.

7.* The **Hermite functions** $h_k(x)$ are defined by the generating identity

$$\sum_{k=0}^{\infty} h_k(x) \frac{t^k}{k!} = e^{-(x^2/2-2tx+t^2)}.$$

(a) Show that an alternate definition of the Hermite functions is given by the formula

$$h_k(x) = (-1)^k e^{x^2/2} \left(\frac{d}{dx}\right)^k e^{-x^2}.$$

[Hint: Write $e^{-(x^2/2-2tx+t^2)} = e^{x^2/2}e^{-(x-t)^2}$ and use Taylor's formula.] Conclude from the above expression that each $h_k(x)$ is of the form $P_k(x)e^{-x^2/2}$, where P_k is a polynomial of degree k. In particular, the Hermite functions belong to the Schwartz space and $h_0(x) = e^{-x^2/2}$, $h_1(x) = 2xe^{-x^2/2}$.

(b) Prove that the family $\{h_k\}_{k=0}^{\infty}$ is complete in the sense that if f is a Schwartz function, and

$$(f, h_k) = \int_{-\infty}^{\infty} f(x)h_k(x)\,dx = 0 \qquad \text{for all } k \geq 0,$$

then $f = 0$. [Hint: Use Exercise 8.]

(c) Define $h_k^*(x) = h_k((2\pi)^{1/2}x)$. Then

$$\widehat{h_k^*}(\xi) = (-i)^k h_k^*(\xi).$$

Therefore, each h_k^* is an eigenfunction for the Fourier transform.

(d) Show that h_k is an eigenfunction for the operator defined in Exercise 23, and in fact, prove that

$$Lh_k = (2k+1)h_k.$$

In particular, we conclude that the functions h_k are mutually orthogonal for the L^2 inner product on the Schwartz space.

(e) Finally, show that $\int_{-\infty}^{\infty} [h_k(x)]^2\,dx = \pi^{1/2}2^k k!$. [Hint: Square the generating relation.]

8.* To refine the results in Chapter 4, and to prove that

$$f_\alpha(x) = \sum_{n=0}^{\infty} 2^{-n\alpha} e^{2\pi i 2^n x}$$

is nowhere differentiable even in the case $\alpha = 1$, we need to consider a variant of the delayed means $\triangle_N$, which in turn will be analyzed by the Poisson summation formula.

(a) Fix an indefinitely differentiable function Φ satisfying

$$\Phi(\xi) = \begin{cases} 1 & \text{when } |\xi| \leq 1, \\ 0 & \text{when } |\xi| \geq 2. \end{cases}$$

By the Fourier inversion formula, there exists $\varphi \in \mathcal{S}$ so that $\hat{\varphi}(\xi) = \Phi(\xi)$. Let $\varphi_N(x) = N\varphi(Nx)$ so that $\widehat{\varphi_N}(\xi) = \Phi(\xi/N)$. Finally, set

$$\tilde{\triangle}_N(x) = \sum_{n=-\infty}^{\infty} \varphi_N(x+n).$$

Observe by the Poisson summation formula that $\tilde{\triangle}_N(x) = \sum_{n=-\infty}^{\infty} \Phi(n/N) e^{2\pi i n x}$, thus $\tilde{\triangle}_N$ is a trigonometric polynomial of degree $\leq 2N$, with terms whose coefficients are 1 when $|n| \leq N$. Let

$$\tilde{\triangle}_N(f) = f * \tilde{\triangle}_N.$$

Note that

$$S_N(f_\alpha) = \tilde{\triangle}_{N'}(f_\alpha)$$

where N' is the largest integer of the form 2^k with $N' \leq N$.

(b) If we set $\tilde{\triangle}_N(x) = \varphi_N(x) + E_N(x)$ where

$$E_N(x) = \sum_{|n| \geq 1} \varphi_N(x+n),$$

then one sees that:

(i) $\sup_{|x| \leq 1/2} |E_N'(x)| \to 0$ as $N \to \infty$.

(ii) $|\tilde{\triangle}_N'(x)| \leq cN^2$.

(iii) $|\tilde{\triangle}_N'(x)| \leq c/(N|x|^3)$, for $|x| \leq 1/2$.

Moreover, $\int_{|x| \leq 1/2} \tilde{\triangle}_N'(x)\, dx = 0$, and $-\int_{|x| \leq 1/2} x \tilde{\triangle}_N'(x)\, dx \to 1$ as $N \to \infty$.

(c) The above estimates imply that if $f'(x_0)$ exists, then

$$(f * \tilde{\triangle}_N')(x_0 + h_N) \to f'(x_0) \qquad \text{as } N \to \infty,$$

whenever $|h_N| \leq C/N$. Then, conclude that both the real and imaginary parts of f_1 are nowhere differentiable, as in the proof given in Section 3, Chapter 4.

6 The Fourier Transform on $\mathbb{R}^d$

> It occurred to me that in order to improve treatment planning one had to know the distribution of the attenuation coefficient of tissues in the body. This information would be useful for diagnostic purposes and would constitute a tomogram or series of tomograms.
>
> It was immediately evident that the problem was a mathematical one. If a fine beam of gamma rays of intensity I_0 is incident on the body and the emerging density is I, then the measurable quantity g equals $\log(I_0/I) = \int_L f\,ds$, where f is the variable absorption coefficient along the line L. Hence if f is a function of two dimensions, and g is known for all lines intersecting the body, the question is, can f be determined if g is known?
>
> Fourteen years would elapse before I learned that Radon had solved this problem in 1917.
>
> *A. M. Cormack,* 1979

The previous chapter introduced the theory of the Fourier transform on $\mathbb{R}$ and illustrated some of its applications to partial differential equations. Here, our aim is to present an analogous theory for functions of several variables.

After a brief review of some relevant notions in $\mathbb{R}^d$, we begin with some general facts about the Fourier transform on the Schwartz space $\mathcal{S}(\mathbb{R}^d)$. Fortunately, the main ideas and techniques have already been considered in the one-dimensional case. In fact, with the appropriate notation, the statements (and proofs) of the key theorems, such as the Fourier inversion and Plancherel formulas, remain unchanged.

Next, we highlight the connection to some higher dimensional problems in mathematical physics, and in particular we investigate the wave equation in d dimensions, with a detailed analysis in the cases $d = 3$ and $d = 2$. At this stage, we discover a rich interplay between the Fourier transform and rotational symmetry, that arises only in $\mathbb{R}^d$ when $d \geq 2$.

Finally, the chapter ends with a discussion of the Radon transform. This topic is of substantial interest in its own right, but in addition has significant relevance in its application to the use of X-ray scans as well

as to other parts of mathematics.

1 Preliminaries

The setting in this chapter will be $\mathbb{R}^d$, the vector space[1] of all d-tuples of real numbers $(x_1, \ldots, x_d)$ with $x_i \in \mathbb{R}$. Addition of vectors is componentwise, and so is multiplication by real scalars. Given $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ we define

$$|x| = (x_1^2 + \cdots + x_d^2)^{1/2},$$

so that $|x|$ is simply the length of the vector x in the usual Euclidean norm. In fact, we equip $\mathbb{R}^d$ with the standard inner product defined by

$$x \cdot y = x_1 y_1 + \cdots + x_d y_d,$$

so that $|x|^2 = x \cdot x$. We use the notation $x \cdot y$ in place of (x, y) of Chapter 3.

Given a d-tuple $\alpha = (\alpha_1, \ldots, \alpha_d)$ of non-negative integers (sometimes called a **multi-index**), the monomial x^α is defined by

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}.$$

Similarly, we define the differential operator $(\partial/\partial x)^\alpha$ by

$$\left(\frac{\partial}{\partial x}\right)^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}},$$

where $|\alpha| = \alpha_1 + \cdots + \alpha_d$ is the order of the multi-index α.

1.1 Symmetries

Analysis in $\mathbb{R}^d$, and in particular the theory of the Fourier transform, is shaped by three important groups of symmetries of the underlying space:

(i) Translations

(ii) Dilations

(iii) Rotations

[1] See Chapter 3 for a brief review of vector spaces and inner products. Here we find it convenient to use lower case letters such as x (as opposed to X) to designate points in $\mathbb{R}^d$. Also, we use $|\cdot|$ instead of $\|\cdot\|$ to denote the Euclidean norm.

We have seen that translations $x \mapsto x + h$, with $h \in \mathbb{R}^d$ fixed, and dilations $x \mapsto \delta x$, with $\delta > 0$, play an important role in the one-dimensional theory. In $\mathbb{R}$, the only two rotations are the identity and multiplication by -1. However, in $\mathbb{R}^d$ with $d \geq 2$ there are more rotations, and the understanding of the interaction between the Fourier transform and rotations leads to fruitful insights regarding spherical symmetries.

A **rotation** in $\mathbb{R}^d$ is a linear transformation $R : \mathbb{R}^d \to \mathbb{R}^d$ which preserves the inner product. In other words,

- $R(ax + by) = aR(x) + bR(y)$ for all $x, y \in \mathbb{R}^d$ and $a, b \in \mathbb{R}$.
- $R(x) \cdot R(y) = x \cdot y$ for all $x, y \in \mathbb{R}^d$.

Equivalently, this last condition can be replaced by $|R(x)| = |x|$ for all $x \in \mathbb{R}^d$, or $R^t = R^{-1}$ where R^t and R^{-1} denote the transpose and inverse of R, respectively.[2] In particular, we have $\det(R) = \pm 1$, where $\det(R)$ is the determinant of R. If $\det(R) = 1$ we say that R is a **proper rotation**; otherwise, we say that R is an **improper rotation**.

EXAMPLE 1. On the real line $\mathbb{R}$, there are two rotations: the identity which is proper, and multiplication by -1 which is improper.

EXAMPLE 2. The rotations in the plane $\mathbb{R}^2$ can be described in terms of complex numbers. We identify $\mathbb{R}^2$ with $\mathbb{C}$ by assigning the point (x, y) to the complex number $z = x + iy$. Under this identification, all proper rotations are of the form $z \mapsto ze^{i\varphi}$ for some $\varphi \in \mathbb{R}$, and all improper rotations are of the form $z \mapsto \overline{z}e^{i\varphi}$ for some $\varphi \in \mathbb{R}$ (here, $\overline{z} = x - iy$ denotes the complex conjugate of z). See Exercise 1 for the argument leading to this result.

EXAMPLE 3. Euler gave the following very simple geometric description of rotations in $\mathbb{R}^3$. Given a proper rotation R, there exists a unit vector γ so that:

(i) R fixes γ, that is, $R(\gamma) = \gamma$.

(ii) If $\mathcal{P}$ denotes the plane passing through the origin and perpendicular to γ, then $R : \mathcal{P} \to \mathcal{P}$, and the restriction of R to $\mathcal{P}$ is a rotation in $\mathbb{R}^2$.

[2]Recall that the transpose of a linear operator $A : \mathbb{R}^d \to \mathbb{R}^d$ is the linear operator $B : \mathbb{R}^d \to \mathbb{R}^d$ which satisfies $A(x) \cdot y = x \cdot B(y)$ for all $x, y \in \mathbb{R}^d$. We write $B = A^t$. The inverse of A (when it exists) is the linear operator $C : \mathbb{R}^d \to \mathbb{R}^d$ with $A \circ C = C \circ A = I$ (where I is the identity), and we write $C = A^{-1}$.

Geometrically, the vector γ gives the direction of the axis of rotation. A proof of this fact is given in Exercise 2. Finally, if R is improper, then $-R$ is proper (since in $\mathbb{R}^3$ $\det(-R) = -\det(R)$), so R is the composition of a proper rotation and a symmetry with respect to the origin.

EXAMPLE 4. Given two *orthonormal* bases $\{e_1, \ldots, e_d\}$ and $\{e'_1, \ldots, e'_d\}$ in $\mathbb{R}^d$, we can define a rotation R by letting $R(e_i) = e'_i$ for $i = 1, \ldots, d$. Conversely, if R is a rotation and $\{e_1, \ldots, e_d\}$ is an orthonormal basis, then $\{e'_1, \ldots, e'_d\}$, where $e'_j = R(e_j)$, is another orthonormal basis.

1.2 Integration on $\mathbb{R}^d$

Since we shall be dealing with functions on $\mathbb{R}^d$, we will need to discuss some aspects of integration of such functions. A more detailed review of integration on $\mathbb{R}^d$ is given in the appendix.

A continuous complex-valued function f on $\mathbb{R}^d$ is said to be **rapidly decreasing** if for every multi-index α the function $|x^\alpha f(x)|$ is bounded. Equivalently, a continuous function is of rapid decrease if

$$\sup_{x\in\mathbb{R}^d} |x|^k \, |f(x)| < \infty \qquad \text{for every } k = 0, 1, 2, \ldots.$$

Given a function of rapid decrease, we define

$$\int_{\mathbb{R}^d} f(x)\, dx = \lim_{N\to\infty} \int_{Q_N} f(x)\, dx,$$

where Q_N denotes the closed cube centered at the origin, with sides of length N parallel to the coordinate axis, that is,

$$Q_N = \{x \in \mathbb{R}^d : \ |x_i| \le N/2 \quad \text{for } i = 1, \ldots, d\}.$$

The integral over Q_N is a multiple integral in the usual sense of Riemann integration. That the limit exists follows from the fact that the integrals $I_N = \int_{Q_N} f(x)\, dx$ form a Cauchy sequence as N tends to infinity.

Two observations are in order. First, we may replace the square Q_N by the ball $B_N = \{x \in \mathbb{R}^d : |x| \le N\}$ without changing the definition. Second, we do not need the full force of rapid decrease to show that the limit exists. In fact it suffices to assume that f is continuous and

$$\sup_{x\in\mathbb{R}^d} |x|^{d+\epsilon} \, |f(x)| < \infty \qquad \text{for some } \epsilon > 0.$$

For example, functions of moderate decrease on $\mathbb{R}$ correspond to $\epsilon = 1$. In keeping with this we define functions of **moderate decrease** on $\mathbb{R}^d$ as those that are continuous and satisfy the above inequality with $\epsilon = 1$.

The interaction of integration with the three important groups of symmetries is as follows: if f is of moderate decrease, then

(i) $\displaystyle\int_{\mathbb{R}^d} f(x+h)\,dx = \int_{\mathbb{R}^d} f(x)\,dx$ for all $h \in \mathbb{R}^d$,

(ii) $\displaystyle\delta^d \int_{\mathbb{R}^d} f(\delta x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx$ for all $\delta > 0$,

(iii) $\displaystyle\int_{\mathbb{R}^d} f(R(x))\,dx = \int_{\mathbb{R}^d} f(x)\,dx$ for every rotation R.

Polar coordinates

It will be convenient to introduce polar coordinates in $\mathbb{R}^d$ and find the corresponding integration formula. We begin with two examples which correspond to the case $d = 2$ and $d = 3$. (A more elaborate discussion applying to all d is contained in the appendix.)

EXAMPLE 1. In $\mathbb{R}^2$, polar coordinates are given by (r, θ) with $r \geq 0$ and $0 \leq \theta < 2\pi$. The Jacobian of the change of variables is equal to r, so that

$$\int_{\mathbb{R}^2} f(x)\,dx = \int_0^{2\pi}\int_0^{\infty} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta.$$

Now we may write a point on the unit circle S^1 as $\gamma = (\cos\theta, \sin\theta)$, and given a function g on the circle, we define its integral over S^1 by

$$\int_{S^1} g(\gamma)\,d\sigma(\gamma) = \int_0^{2\pi} g(\cos\theta, \sin\theta)\,d\theta.$$

With this notation we then have

$$\int_{\mathbb{R}^2} f(x)\,dx = \int_{S^1}\int_0^{\infty} f(r\gamma)\,r\,dr\,d\sigma(\gamma).$$

EXAMPLE 2. In $\mathbb{R}^3$ one uses spherical coordinates given by

$$\begin{cases} x_1 &= r\sin\theta\cos\varphi, \\ x_2 &= r\sin\theta\sin\varphi, \\ x_3 &= r\cos\theta, \end{cases}$$

where $0 < r$, $0 \leq \theta \leq \pi$ and $0 \leq \varphi \leq 2\pi$. The Jacobian of the change of variables is $r^2 \sin\theta$ so that

$$\int_{\mathbb{R}^3} f(x)\,dx =$$
$$\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta) r^2 dr \sin\theta\, d\theta\, d\varphi.$$

If g is a function on the unit sphere $S^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$, and $\gamma = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, we define the surface element $d\sigma(\gamma)$ by

$$\int_{S^2} g(\gamma)\,d\sigma(\gamma) = \int_0^{2\pi}\int_0^{\pi} g(\gamma)\sin\theta\,d\theta\,d\varphi.$$

As a result,

$$\int_{\mathbb{R}^3} f(x)\,dx = \int_{S^2}\int_0^{\infty} f(r\gamma)\,r^2\,dr\,d\sigma(\gamma).$$

In general, it is possible to write any point in $\mathbb{R}^d - \{0\}$ uniquely as

$$x = r\gamma$$

where γ lies on the unit sphere $S^{d-1} \subset \mathbb{R}^d$ and $r > 0$. Indeed, take $r = |x|$ and $\gamma = x/|x|$. Thus one may proceed as in the cases $d = 2$ or $d = 3$ to define spherical coordinates. The formula we shall use is

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_{S^{d-1}}\int_0^{\infty} f(r\gamma)\,r^{d-1}\,dr\,d\sigma(\gamma),$$

whenever f is of moderate decrease. Here $d\sigma(\gamma)$ denotes the surface element on the sphere S^{d-1} obtained from the spherical coordinates.

2 Elementary theory of the Fourier transform

The **Schwartz space** $\mathcal{S}(\mathbb{R}^d)$ (sometimes abbreviated as $\mathcal{S}$) consists of all indefinitely differentiable functions f on $\mathbb{R}^d$ such that

$$\sup_{x\in\mathbb{R}^d}\left|x^{\alpha}\left(\frac{\partial}{\partial x}\right)^{\beta} f(x)\right| < \infty,$$

for every multi-index α and β. In other words, f and all its derivatives are required to be rapidly decreasing.

EXAMPLE 1. An example of a function in $\mathcal{S}(\mathbb{R}^d)$ is the d-dimensional Gaussian given by $e^{-\pi|x|^2}$. The theory in Chapter 5 already made clear the central role played by this function in the case $d = 1$.

The **Fourier transform** of a Schwartz function f is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}\,dx, \qquad \text{for } \xi \in \mathbb{R}^d.$$

Note the resemblance with the formula in one-dimension, except that we are now integrating on $\mathbb{R}^d$, and the product of x and ξ is replaced by the inner product of the two vectors.

We now list some simple properties of the Fourier transform. In the next proposition the arrow indicates that we have taken the Fourier transform, so $F(x) \longrightarrow G(\xi)$ means that $G(\xi) = \hat{F}(\xi)$.

Proposition 2.1 *Let* $f \in \mathcal{S}(\mathbb{R}^d)$.

(i) $f(x+h) \longrightarrow \hat{f}(\xi)e^{2\pi i\xi\cdot h}$ *whenever* $h \in \mathbb{R}^d$.

(ii) $f(x)e^{-2\pi i x h} \longrightarrow \hat{f}(\xi + h)$ *whenever* $h \in \mathbb{R}^d$.

(iii) $f(\delta x) \longrightarrow \delta^{-d}\hat{f}(\delta^{-1}\xi)$ *whenever* $\delta > 0$.

(iv) $\left(\frac{\partial}{\partial x}\right)^\alpha f(x) \longrightarrow (2\pi i\xi)^\alpha \hat{f}(\xi)$.

(v) $(-2\pi i x)^\alpha f(x) \longrightarrow \left(\frac{\partial}{\partial \xi}\right)^\alpha \hat{f}(\xi)$.

(vi) $f(Rx) \longrightarrow \hat{f}(R\xi)$ *whenever* R *is a rotation.*

The first five properties are proved in the same way as in the one-dimensional case. To verify the last property, simply change variables $y = Rx$ in the integral. Then, recall that $|\det(R)| = 1$, and $R^{-1}y\cdot\xi = y\cdot R\xi$, because R is a rotation.

Properties (iv) and (v) in the proposition show that, up to factors of $2\pi i$, the Fourier transform interchanges differentiation and multiplication by monomials. This motivates the definition of the Schwartz space and leads to the next corollary.

Corollary 2.2 *The Fourier transform maps* $\mathcal{S}(\mathbb{R}^d)$ *to itself.*

At this point we disgress to observe a simple fact concerning the interplay between the Fourier transform and rotations. We say that a

function f is **radial** if it depends only on $|x|$; in other words, f is radial if there is a function $f_0(u)$, defined for $u \geq 0$, such that $f(x) = f_0(|x|)$. We note that f is radial if and only if $f(Rx) = f(x)$ for every rotation R. In one direction, this is obvious since $|Rx| = |x|$. Conversely, suppose that $f(Rx) = f(x)$, for all rotations R. Now define f_0 by

$$f_0(u) = \begin{cases} f(0) & \text{if } u = 0, \\ f(x) & \text{if } |x| = u. \end{cases}$$

Note that f_0 is well defined, since if x and x' are points with $|x| = |x'|$ there is always a rotation R so that $x' = Rx$.

Corollary 2.3 *The Fourier transform of a radial function is radial.*

This follows at once from property (vi) in the last proposition. Indeed, the condition $f(Rx) = f(x)$ for all R implies that $\hat{f}(R\xi) = \hat{f}(\xi)$ for all R, thus $\hat{f}$ is radial whenever f is.

An example of a radial function in $\mathbb{R}^d$ is the Gaussian $e^{-\pi|x|^2}$. Also, we observe that when $d = 1$, the radial functions are precisely the even functions, that is, those for which $f(x) = f(-x)$.

After these preliminaries, we retrace the steps taken in the previous chapter to obtain the Fourier inversion formula and Plancherel theorem for $\mathbb{R}^d$.

Theorem 2.4 *Suppose $f \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i x \cdot \xi}\, d\xi.$$

Moreover

$$\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\, d\xi = \int_{\mathbb{R}^d} |f(x)|^2\, dx.$$

The proof proceeds in the following stages.

Step 1. The Fourier transform of $e^{-\pi|x|^2}$ is $e^{-\pi|\xi|^2}$. To prove this, notice that the properties of the exponential functions imply that

$$e^{-\pi|x|^2} = e^{-\pi x_1^2} \cdots e^{-\pi x_d^2} \quad \text{and} \quad e^{-2\pi i x \cdot \xi} = e^{-2\pi i x_1 \cdot \xi_1} \cdots e^{-2\pi i x_d \cdot \xi_d},$$

so that the integrand in the Fourier transform is a product of d functions, each depending on the variable x_j $(1 \leq j \leq d)$ only. Thus the assertion

follows by writing the integral over $\mathbb{R}^d$ as a series of repeated integrals, each taken over $\mathbb{R}$. For example, when $d = 2$,

$$\begin{aligned}\int_{\mathbb{R}^2} e^{-\pi|x|^2} e^{-2\pi i x\cdot\xi}\, dx &= \int_{\mathbb{R}} e^{-\pi x_2^2} e^{-2\pi i x_2\cdot\xi_2} \left(\int_{\mathbb{R}} e^{-\pi x_1^2} e^{-2\pi i x_1\cdot\xi_1} dx_1\right) dx_2 \\ &= \int_{\mathbb{R}} e^{-\pi x_2^2} e^{-2\pi i x_2\cdot\xi_2} e^{-\pi\xi_1^2}\, dx_2 \\ &= e^{-\pi\xi_1^2} e^{-\pi\xi_2^2} \\ &= e^{-\pi|\xi|^2}.\end{aligned}$$

As a consequence of Proposition 2.1, applied with $\delta^{1/2}$ instead of δ, we find that $\widehat{(e^{-\pi\delta|x|^2})} = \delta^{-d/2} e^{-\pi|\xi|^2/\delta}$.

Step 2. The family $K_\delta(x) = \delta^{-d/2} e^{-\pi|x|^2/\delta}$ is a family of good kernels in $\mathbb{R}^d$. By this we mean that

(i) $\displaystyle\int_{\mathbb{R}^d} K_\delta(x)\, dx = 1$,

(ii) $\displaystyle\int_{\mathbb{R}^d} |K_\delta(x)|\, dx \leq M$ (in fact $K_\delta(x) \geq 0$),

(iii) For every $\eta > 0$, $\displaystyle\int_{|x|\geq\eta} |K_\delta(x)|\, dx \to 0$ as $\delta \to 0$.

The proofs of these assertions are almost identical to the case $d = 1$. As a result

$$\int_{\mathbb{R}^d} K_\delta(x) F(x)\, dx \to F(0) \qquad \text{as } \delta \to 0$$

when F is a Schwartz function, or more generally when F is bounded and continuous at the origin.

Step 3. The multiplication formula

$$\int_{\mathbb{R}^d} f(x)\hat{g}(x)\, dx = \int_{\mathbb{R}^d} \hat{f}(y) g(y)\, dy$$

holds whenever f and g are in $\mathcal{S}$. The proof requires the evaluation of the integral of $f(x)g(y)e^{-2\pi i x\cdot y}$ over $(x, y) \in \mathbb{R}^{2d} = \mathbb{R}^d \times \mathbb{R}^d$ as a repeated integral, with each separate integration taken over $\mathbb{R}^d$. The justification is similar to that in the proof of Proposition 1.8 in the previous chapter. (See the appendix.)

The Fourier inversion is then a simple consequence of the multiplication formula and the family of good kernels K_δ, as in Chapter 5. It also

follows that the Fourier transform $\mathcal{F}$ is a bijective map of $\mathcal{S}(\mathbb{R}^d)$ to itself, whose inverse is

$$\mathcal{F}^*(g)(x) = \int_{\mathbb{R}^d} g(\xi)e^{2\pi i x\cdot\xi}\,d\xi.$$

Step 4. Next we turn to the convolution, defined by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy, \qquad f, g \in \mathcal{S}.$$

We have that $f * g \in \mathcal{S}(\mathbb{R}^d)$, $f * g = g * f$, and $\widehat{(f * g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$. The argument is similar to that in one-dimension. The calculation of the Fourier transform of $f * g$ involves an integration of $f(y)g(x-y)e^{-2\pi i x\cdot\xi}$ (over $\mathbb{R}^{2d} = \mathbb{R}^d \times \mathbb{R}^d$) expressed as a repeated integral.

Then, following the same argument in the previous chapter, we obtain the d-dimensional Plancherel formula, thereby concluding the proof of Theorem 2.4.

3 The wave equation in $\mathbb{R}^d \times \mathbb{R}$

Our next goal is to apply what we have learned about the Fourier transform to the study of the wave equation. Here, we once again simplify matters by restricting ourselves to functions in the Schwartz class $\mathcal{S}$. We note that in any further analysis of the wave equation it is important to allow functions that have much more general behavior, and in particular that may be discontinuous. However, what we lose in generality by only considering Schwartz functions, we gain in transparency. Our study in this restricted context will allow us to explain certain basic ideas in their simplest form.

3.1 Solution in terms of Fourier transforms

The motion of a vibrating string satisfies the equation

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2},$$

which we referred to as the one-dimensional wave equation.

A natural generalization of this equation to d space variables is

$$\frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_d^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}. \tag{1}$$

In fact, it is known that in the case $d = 3$, this equation determines the behavior of electromagnetic waves in vacuum (with c = speed of light).

Also, this equation describes the propagation of sound waves. Thus (1) is called the **d-dimensional wave equation**.

Our first observation is that we may assume $c = 1$, since we can rescale the variable t if necessary. Also, if we define the **Laplacian** in d dimensions by

$$\triangle = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_d^2},$$

then the wave equation can be rewritten as

$$\triangle u = \frac{\partial^2 u}{\partial t^2}. \tag{2}$$

The goal of this section is to find a solution to this equation, subject to the initial conditions

$$u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x),$$

where $f, g \in \mathcal{S}(\mathbb{R}^d)$. This is called the **Cauchy problem** for the wave equation.

Before solving this problem, we note that while we think of the variable t as time, we do not restrict ourselves to $t > 0$. As we will see, the solution we obtain makes sense for all $t \in \mathbb{R}$. This is a manifestation of the fact that the wave equation can be reversed in time (unlike the heat equation).

A formula for the solution of our problem is given in the next theorem. The heuristic argument which leads to this formula is important since, as we have already seen, it applies to some other boundary value problems as well.

Suppose u solves the Cauchy problem for the wave equation. The technique employed consists of taking the Fourier transform of the equation and of the initial conditions, with respect to the space variables $x_1, \ldots, x_d$. This reduces the problem to an ordinary differential equation in the time variable. Indeed, recalling that differentiation with respect to x_j becomes multiplication by $2\pi i \xi_j$, and the differentiation with respect to t commutes with the Fourier transform in the space variables, we find that (2) becomes

$$-4\pi^2 |\xi|^2 \hat{u}(\xi, t) = \frac{\partial^2 \hat{u}}{\partial t^2}(\xi, t).$$

For each fixed $\xi \in \mathbb{R}^d$, this is an ordinary differential equation in t whose solution is given by

$$\hat{u}(\xi, t) = A(\xi) \cos(2\pi|\xi|t) + B(\xi) \sin(2\pi|\xi|t),$$

where for each ξ, $A(\xi)$ and $B(\xi)$ are unknown constants to be determined by the initial conditions. In fact, taking the Fourier transform (in x) of the initial conditions yields

$$\hat{u}(\xi, 0) = \hat{f}(\xi) \quad \text{and} \quad \frac{\partial \hat{u}}{\partial t}(\xi, 0) = \hat{g}(\xi).$$

We may now solve for $A(\xi)$ and $B(\xi)$ to obtain

$$A(\xi) = \hat{f}(\xi) \quad \text{and} \quad 2\pi|\xi|B(\xi) = \hat{g}(\xi).$$

Therefore, we find that

$$\hat{u}(\xi, t) = \hat{f}(\xi)\cos(2\pi|\xi|t) + \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|},$$

and the solution u is given by taking the inverse Fourier transform in the ξ variables. This formal derivation then leads to a precise existence theorem for our problem.

Theorem 3.1 *A solution of the Cauchy problem for the wave equation is*

$$u(x,t) = \int_{\mathbb{R}^d} \left[\hat{f}(\xi)\cos(2\pi|\xi|t) + \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\right] e^{2\pi i x\cdot\xi}\, d\xi. \tag{3}$$

Proof. We first verify that u solves the wave equation. This is straightforward once we note that we can differentiate in x and t under the integral sign (because f and g are both Schwartz functions) and therefore u is at least C^2. On the one hand we differentiate the exponential with respect to the x variables to get

$$\triangle u(x,t) = \int_{\mathbb{R}^d} \left[\hat{f}(\xi)\cos(2\pi|\xi|t) + \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\right] (-4\pi^2|\xi|^2) e^{2\pi i x\cdot\xi}\, d\xi,$$

while on the other hand we differentiate the terms in brackets with respect to t twice to get

$$\frac{\partial^2 u}{\partial t^2}(x,t) =$$

$$\int_{\mathbb{R}^d} \left[-4\pi^2|\xi|^2\hat{f}(\xi)\cos(2\pi|\xi|t) - 4\pi^2|\xi|^2\hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\right] e^{2\pi i x\cdot\xi}\, d\xi.$$

This shows that u solves equation (2). Setting $t = 0$ we get

$$u(x,0) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i x\cdot\xi}\, d\xi = f(x)$$

by the Fourier inversion theorem. Finally, differentiating once with respect to t, setting $t = 0$, and using the Fourier inversion shows that

$$\frac{\partial u}{\partial t}(x, 0) = g(x).$$

Thus u also verifies the initial conditions, and the proof of the theorem is complete.

As the reader will note, both $\hat{f}(\xi)\cos(2\pi|\xi|t)$ and $\hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}$ are functions in $\mathcal{S}$, assuming as we do that f and g are in $\mathcal{S}$. This is because both $\cos u$ and $(\sin u)/u$ are even functions that are indefinitely differentiable.

Having proved the existence of a solution to the Cauchy problem for the wave equation, we raise the question of uniqueness. Are there solutions to the problem

$$\triangle u = \frac{\partial^2 u}{\partial t^2} \quad \text{subject to} \quad u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x),$$

other than the one given by the formula in the theorem? In fact the answer is, as expected, no. The proof of this fact, which will not be given here (but see Problem 3), can be based on a conservation of energy argument. This is a local counterpart of a global conservation of energy statement which we will now present.

We observed in Exercise 10, Chapter 3, that in the one-dimensional case, the total energy of the vibrating string is conserved in time. The analogue of this fact holds in higher dimensions as well. Define the **energy** of a solution by

$$E(t) = \int_{\mathbb{R}^d} \left|\frac{\partial u}{\partial t}\right|^2 + \left|\frac{\partial u}{\partial x_1}\right|^2 + \cdots + \left|\frac{\partial u}{\partial x_d}\right|^2 dx.$$

Theorem 3.2 *If u is the solution of the wave equation given by formula (3), then $E(t)$ is conserved, that is,*

$$E(t) = E(0), \qquad \textit{for all } t \in \mathbb{R}.$$

The proof requires the following lemma.

Lemma 3.3 *Suppose a and b are complex numbers and α is real. Then*

$$|a\cos\alpha + b\sin\alpha|^2 + |-a\sin\alpha + b\cos\alpha|^2 = |a|^2 + |b|^2.$$

This follows directly because $e_1 = (\cos\alpha, \sin\alpha)$ and $e_2 = (-\sin\alpha, \cos\alpha)$ are a pair of orthonormal vectors, hence with $Z = (a, b) \in \mathbb{C}^2$, we have

$$|Z|^2 = |Z \cdot e_1|^2 + |Z \cdot e_2|^2,$$

where $\cdot$ represents the inner product in $\mathbb{C}^2$.

Now by Plancherel's theorem,

$$\int_{\mathbb{R}^d} \left|\frac{\partial u}{\partial t}\right|^2 dx = \int_{\mathbb{R}^d} \left|-2\pi|\xi|\hat{f}(\xi)\sin(2\pi|\xi|t) + \hat{g}(\xi)\cos(2\pi|\xi|t)\right|^2 d\xi.$$

Similarly,

$$\int_{\mathbb{R}^d} \sum_{j=1}^{d} \left|\frac{\partial u}{\partial x_j}\right|^2 dx = \int_{\mathbb{R}^d} \left|2\pi|\xi|\hat{f}(\xi)\cos(2\pi|\xi|t) + \hat{g}(\xi)\sin(2\pi|\xi|t)\right|^2 d\xi.$$

We now apply the lemma with

$$a = 2\pi|\xi|\hat{f}(\xi), \qquad b = \hat{g}(\xi) \qquad \text{and} \qquad \alpha = 2\pi|\xi|t.$$

The result is that

$$\begin{aligned} E(t) &= \int_{\mathbb{R}^d} \left|\frac{\partial u}{\partial t}\right|^2 + \left|\frac{\partial u}{\partial x_1}\right|^2 + \cdots + \left|\frac{\partial u}{\partial x_d}\right|^2 dx \\ &= \int_{\mathbb{R}^d} (4\pi^2|\xi|^2|\hat{f}(\xi)|^2 + |\hat{g}(\xi)|^2)\, d\xi, \end{aligned}$$

which is clearly independent of t. Thus Theorem 3.2 is proved.

The drawback with formula (3), which does give the solution of the wave equation, is that it is quite indirect, involving the calculation of the Fourier transforms of f and g, and then a further inverse Fourier transform. However, for every dimension d there is a more explicit formula. This formula is very simple when $d = 1$ and a little less so when $d = 3$. More generally, the formula is "elementary" whenever d is odd, and more complicated when d is even (see Problems 4 and 5).

In what follows we consider the cases $d = 1$, $d = 3$, and $d = 2$, which together give a picture of the general situation. Recall that in Chapter 1, when discussing the wave equation over the interval $[0, L]$, we found that the solution is given by d'Alembert's formula

$$u(x,t) = \frac{f(x+t) + f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(y)\, dy. \tag{4}$$

with the interpretation that both f and g are extended outside $[0, L]$ by making them *odd* in $[-L, L]$, and periodic on the real line, with period $2L$. The same formula (4) holds for the solution of the wave equation when $d = 1$ and when the initial data are functions in $\mathcal{S}(\mathbb{R})$. In fact, this follows directly from (3) if we note that

$$\cos(2\pi|\xi|t) = \frac{1}{2}(e^{2\pi i|\xi|t} + e^{-2\pi i|\xi|t})$$

and

$$\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|} = \frac{1}{4\pi i|\xi|}(e^{2\pi i|\xi|t} - e^{-2\pi i|\xi|t}).$$

Finally, we note that the two terms that appear in d'Alembert's formula (4) consist of appropriate averages. Indeed, the first term is precisely the average of f over the two points that are the boundary of the interval $[x - t, x + t]$; the second term is, up to a factor of t, the mean value of g over this interval, that is, $(1/2t)\int_{x-t}^{x+t} g(y)\, dy$. This suggests a generalization to higher dimensions, where we might expect to write the solution of our problem as averages of the initial data. This is in fact the case, and we now treat in detail the particular situation $d = 3$.

3.2 The wave equation in $\mathbb{R}^3 \times \mathbb{R}$

If S^2 denotes the unit sphere in $\mathbb{R}^3$, we define the **spherical mean** of the function f over the sphere of radius t centered at x by

$$M_t(f)(x) = \frac{1}{4\pi}\int_{S^2} f(x - t\gamma)\, d\sigma(\gamma), \tag{5}$$

where $d\sigma(\gamma)$ is the element of surface area for S^2. Since 4π is the area of the unit sphere, we can interpret $M_t(f)$ as the average value of f over the sphere centered at x of radius t.

Lemma 3.4 *If $f \in \mathcal{S}(\mathbb{R}^3)$ and t is fixed, then $M_t(f) \in \mathcal{S}(\mathbb{R}^3)$. Moreover, $M_t(f)$ is indefinitely differentiable in t, and each t-derivative also belongs to $\mathcal{S}(\mathbb{R}^3)$.*

Proof. Let $F(x) = M_t(f)(x)$. To show that F is rapidly decreasing, start with the inequality $|f(x)| \leq A_N/(1 + |x|^N)$ which holds for every fixed $N \geq 0$. As a simple consequence, whenever t is fixed, we have

$$|f(x - \gamma t)| \leq A'_N/(1 + |x|^N) \quad \text{for all } \gamma \in S^2.$$

To see this consider separately the cases when $|x| \leq 2|t|$, and $|x| > 2|t|$. Therefore, by integration

$$|F(x)| \leq A'_N/(1+|x|^N),$$

and since this holds for every N, the function F is rapidly decreasing. One next observes that F is indefinitely differentiable, and

$$\left(\frac{\partial}{\partial x}\right)^{\alpha} F(x) = M_t(f^{(\alpha)})(x) \tag{6}$$

where $f^{(\alpha)}(x) = (\partial/\partial x)^{\alpha} f$. It suffices to prove this when $(\partial/\partial x)^{\alpha} = \partial/\partial x_k$, and then proceed by induction to get the general case. Furthermore, it is enough to take $k = 1$. Now

$$\frac{F(x_1+h, x_2, x_3) - F(x_1, x_2, x_3)}{h} = \frac{1}{4\pi}\int_{S^2} g_h(\gamma)\, d\sigma(\gamma)$$

where

$$g_h(\gamma) = \frac{f(x + e_1 h - \gamma t) - f(x - \gamma t)}{h},$$

and $e_1 = (1, 0, 0)$. Now, it suffices to observe that $g_h \to \frac{\partial}{\partial x_1} f(x - \gamma t)$ as $h \to 0$ uniformly in γ. As a result, we find that (6) holds, and by the first argument, it follows that $\left(\frac{\partial}{\partial x}\right)^{\alpha} F(x)$ is also rapidly decreasing, hence $F \in \mathcal{S}$. The same argument applies to each t-derivative of $M_t(f)$.

The basic fact about integration on spheres that we shall need is the following Fourier transform formula.

Lemma 3.5 $\displaystyle \frac{1}{4\pi}\int_{S^2} e^{-2\pi i \xi\cdot\gamma}\, d\sigma(\gamma) = \frac{\sin(2\pi|\xi|)}{2\pi|\xi|}.$

This formula, as we shall see in the following section, is connected to the fact that the Fourier transform of a radial function is radial.

Proof. Note that the integral on the left is radial in ξ. Indeed, if R is a rotation then

$$\int_{S^2} e^{-2\pi i R(\xi)\cdot\gamma}\, d\sigma(\gamma) = \int_{S^2} e^{-2\pi i \xi\cdot R^{-1}(\gamma)}\, d\sigma(\gamma) = \int_{S^2} e^{-2\pi i \xi\cdot\gamma}\, d\sigma(\gamma)$$

because we may change variables $\gamma \to R^{-1}(\gamma)$. (For this, see formula (4) in the appendix.) So if $|\xi| = \rho$, it suffices to prove the lemma with

$\xi = (0,0,\rho)$. If $\rho = 0$, the lemma is obvious. If $\rho > 0$, we choose spherical coordinates to find that the left-hand side is equal to

$$\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} e^{-2\pi i\rho\cos\theta}\sin\theta\, d\theta\, d\varphi.$$

The change of variables $u = -\cos\theta$ gives

$$\begin{aligned}\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} e^{-2\pi i\rho\cos\theta}\sin\theta\, d\theta\, d\varphi &= \frac{1}{2}\int_0^{\pi} e^{-2\pi i\rho\cos\theta}\sin\theta\, d\theta \\ &= \frac{1}{2}\int_{-1}^{1} e^{2\pi i\rho u}\, du \\ &= \frac{1}{4\pi i\rho}\left[e^{2\pi i\rho u}\right]_{-1}^{1} \\ &= \frac{\sin(2\pi\rho)}{2\pi\rho},\end{aligned}$$

and the formula is proved.

By the defining formula (5) we may interpret $M_t(f)$ as a convolution of the function f with the element $d\sigma$, and since the Fourier transform interchanges convolutions with products, we are led to believe that $\widehat{M_t(f)}$ is the product of the corresponding Fourier transforms. Indeed, we have the identity

$$\widehat{M_t(f)}(\xi) = \hat{f}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|t}. \tag{7}$$

To see this, write

$$\widehat{M_t(f)}(\xi) = \int_{\mathbb{R}^3} e^{-2\pi i x\cdot\xi}\left(\frac{1}{4\pi}\int_{S^2} f(x-\gamma t)\, d\sigma(\gamma)\right)dx,$$

and note that we may interchange the order of integration and make a simple change of variables to achieve the desired identity.

As a result, we find that the solution of our problem may be expressed by using the spherical means of the initial data.

Theorem 3.6 *The solution when $d = 3$ of the Cauchy problem for the wave equation*

$$\triangle u = \frac{\partial^2 u}{\partial t^2} \quad \textit{subject to} \quad u(x,0) = f(x) \quad \textit{and} \quad \frac{\partial u}{\partial t}(x,0) = g(x)$$

is given by

$$u(x,t) = \frac{\partial}{\partial t}(tM_t(f)(x)) + tM_t(g)(x).$$

Proof. Consider first the problem

$$\triangle u = \frac{\partial^2 u}{\partial t^2} \quad \text{subject to} \quad u(x,0) = 0 \quad \text{and} \quad \frac{\partial u}{\partial t}(x,0) = g(x).$$

Then by Theorem 3.1, we know that its solution u_1 is given by

$$\begin{aligned} u_1(x,t) &= \int_{\mathbb{R}^3} \left[\hat{g}(\xi) \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|} \right] e^{2\pi i x\cdot\xi}\, d\xi \\ &= t \int_{\mathbb{R}^3} \left[\hat{g}(\xi) \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|t} \right] e^{2\pi i x\cdot\xi}\, d\xi \\ &= tM_t(g)(x), \end{aligned}$$

where we have used (7) applied to g, and the Fourier inversion formula.

According to Theorem 3.1 again, the solution to the problem

$$\triangle u = \frac{\partial^2 u}{\partial t^2} \quad \text{subject to} \quad u(x,0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x,0) = 0$$

is given by

$$\begin{aligned} u_2(x,t) &= \int_{\mathbb{R}^3} \left[\hat{f}(\xi) \cos(2\pi|\xi|t) \right] e^{2\pi i x\cdot\xi}\, d\xi \\ &= \frac{\partial}{\partial t} \left(t \int_{\mathbb{R}^3} \left[\hat{f}(\xi) \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|t} \right] e^{2\pi i x\cdot\xi}\, d\xi \right) \\ &= \frac{\partial}{\partial t}(tM_t(f)(x)). \end{aligned}$$

We may now superpose these two solutions to obtain $u = u_1 + u_2$ as the solution of our original problem.

Huygens principle

The solutions to the wave equation in one and three dimensions are given, respectively, by

$$u(x,t) = \frac{f(x+t) + f(x-t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} g(y)\, dy$$

and

$$u(x,t) = \frac{\partial}{\partial t}(tM_t(f)(x)) + tM_t(g)(x).$$

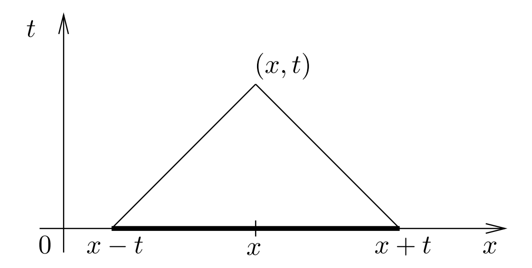


Figure 1. Huygens principle, $d = 1$

We observe that in the one-dimensional problem, the value of the solution at (x, t) depends only on the values of f and g in the interval centered at x of length $2t$, as shown in Figure 1.

If in addition $g = 0$, then the solution depends only on the data at the two boundary points of this interval. In three dimensions, this boundary dependence always holds. More precisely, the solution $u(x, t)$ depends only on the values of f and g in an immediate neighborhood of the sphere centered at x and of radius t. This situation is depicted in Figure 2, where we have drawn the cone originating at (x, t) and with its base the ball centered at x of radius t. This cone is called the **backward light cone** originating at (x, t).

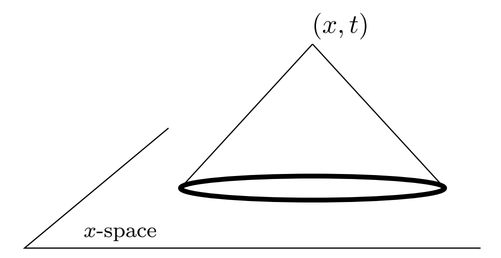


Figure 2. Backward light cone originating at (x, t)

Alternatively, the data at a point x_0 in the plane $t = 0$ influences the solution only on the boundary of a cone originating at x_0, called the **forward light cone** and depicted in Figure 3.

This phenomenon, known as the **Huygens principle**, is immediate from the formulas for u given above.

Another important aspect of the wave equation connected with these

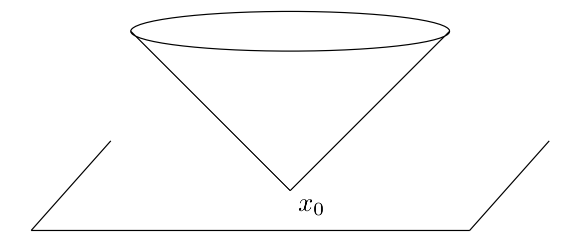


Figure 3. The forward light cone originating at x_0

considerations is that of the **finite speed of propagation**. (In the case where $c = 1$, the speed is 1.) This means that if we have an initial disturbance localized at $x = x_0$, then after a finite time t, its effects will have propagated only inside the ball centered at x_0 of radius $|t|$. To state this precisely, suppose the initial conditions f and g are supported in the ball of radius δ, centered at x_0 (think of δ as small). Then $u(x,t)$ is supported in the ball of radius $|t| + \delta$ centered at x_0. This assertion is clear from the above discussion.

3.3 The wave equation in $\mathbb{R}^2 \times \mathbb{R}$: descent

It is a remarkable fact that the solution of the wave equation in three dimensions leads to a solution of the wave equation in two dimensions. Define the corresponding means by

$$\widetilde{M}_t(F)(x) = \frac{1}{2\pi}\int_{|y|\leq 1} F(x-ty)(1-|y|^2)^{-1/2}\,dy.$$

Theorem 3.7 *A solution of the Cauchy problem for the wave equation in two dimensions with initial data $f, g \in \mathcal{S}(\mathbb{R}^2)$ is given by*

$$u(x,t) = \frac{\partial}{\partial t}(t\widetilde{M}_t(f)(x)) + t\widetilde{M}_t(g)(x). \tag{8}$$

Notice the difference between this case and the case $d = 3$. Here, u at (x,t) depends on f and g in the whole disc (of radius $|t|$ centered at x), and not just on the values of the initial data near the boundary of that disc.

Formally, the identity in the theorem arises as follows. If we start with an initial pair of functions f and g in $\mathcal{S}(\mathbb{R}^2)$, we may consider the corresponding functions $\tilde{f}$ and $\tilde{g}$ on $\mathbb{R}^3$ that are merely extensions of f and g that are constant in the x_3 variable, that is,

$$\tilde{f}(x_1,x_2,x_3) = f(x_1,x_2) \quad \text{and} \quad \tilde{g}(x_1,x_2,x_3) = g(x_1,x_2).$$

Now, if $\tilde{u}$ is the solution (given in the previous section) of the 3-dimensional wave equation with initial data $\tilde{f}$ and $\tilde{g}$, then one can expect that $\tilde{u}$ is also constant in x_3 so that $\tilde{u}$ satisfies the 2-dimensional wave equation. A difficulty with this argument is that $\tilde{f}$ and $\tilde{g}$ are not rapidly decreasing since they are constant in x_3, so that our previous methods do not apply. However, it is easy to modify the argument so as to obtain a proof of Theorem 3.7.

We fix $T > 0$ and consider a function $\eta(x_3)$ that is in $\mathcal{S}(\mathbb{R})$, such that $\eta(x_3) = 1$ if $|x_3| \leq 3T$. The trick is to truncate $\tilde{f}$ and $\tilde{g}$ in the x_3-variable, and consider instead

$$\tilde{f}^\flat(x_1, x_2, x_3) = f(x_1, x_2)\eta(x_3) \quad \text{and} \quad \tilde{g}^\flat(x_1, x_2, x_3) = g(x_1, x_2)\eta(x_3).$$

Now both $\tilde{f}^\flat$ and $\tilde{g}^\flat$ are in $\mathcal{S}(\mathbb{R}^3)$, so Theorem 3.6 provides a solution $\tilde{u}^\flat$ of the wave equation with initial data $\tilde{f}^\flat$ and $\tilde{g}^\flat$. It is easy to see from the formula that $\tilde{u}^\flat(x, t)$ is independent of x_3, whenever $|x_3| \leq T$ and $|t| \leq T$. In particular, if we define $u(x_1, x_2, t) = \tilde{u}^\flat(x_1, x_2, 0, t)$, then u satisfies the 2-dimensional wave equation when $|t| \leq T$. Since T is arbitrary, u is a solution to our problem, and it remains to see why u has the desired form.

By definition of the spherical coordinates, we recall that the integral of a function H over the sphere S^2 is given by

$$\frac{1}{4\pi}\int_{S^2} H(\gamma)\,d\sigma(\gamma) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} H(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\sin\theta\,d\theta\,d\varphi.$$

If H does not depend on the last variable, that is, $H(x_1, x_2, x_3) = h(x_1, x_2)$ for some function h of two variables, then

$$M_t(H)(x_1, x_2, 0) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} h(x_1 - t\sin\theta\cos\varphi, x_2 - t\sin\theta\sin\varphi)\sin\theta\,d\theta\,d\varphi.$$

To calculate this last integral, we split the θ-integral from 0 to $\pi/2$ and then $\pi/2$ to π. By making the change of variables $r = \sin\theta$, we find, after a final change to polar coordinates, that

$$\begin{aligned} M_t(H)(x_1, x_2, 0) &= \frac{1}{2\pi}\int_{|y|\leq 1} h(x - ty)(1 - |y|^2)^{-1/2}\,dy \\ &= \widetilde{M}_t(h)(x_1, x_2). \end{aligned}$$

Applying this to $H = \tilde{f}^\flat$, $h = f$, and $H = \tilde{g}^\flat$, $h = g$, we find that u is given by the formula (8), and the proof of Theorem 3.7 is complete.

Remark. In the case of general d, the solution of the wave equation shares many of the properties we have discussed in the special cases $d = 1, 2$, and 3.

- At a given time t, the initial data at a point x only affects the solution u in a specific region. When $d > 1$ is odd, the data influences only the points on the boundary of the forward light cone originating at x, while when $d = 1$ or d is even, it affects all points of the forward light cone. Alternatively, the solution at a point (x, t) depends only on the data at the base of the backward light cone originating at (x, t). In fact, when $d > 1$ is odd, only the data in an immediate neighborhood of the boundary of the base will influence $u(x, t)$.

- Waves propagate with finite speed: if the initial data is supported in a bounded set, then the support of the solution u spreads with velocity 1 (or more generally c, if the wave equation is not normalized).

We can illustrate some of these facts by the following observation about the different behavior of the propagation of waves in three and two dimensions. Since the propagation of light is governed by the three-dimensional wave equation, if at $t = 0$ a light flashes at the origin, the following happens: any observer will see the flash (after a finite amount of time) only for an instant. In contrast, consider what happens in two dimensions. If we drop a stone in a lake, any point on the surface will begin (after some time) to undulate; although the amplitude of the oscillations will decrease over time, the undulations will continue (in principle) indefinitely.

The difference in character of the formulas for the solutions of the wave equation when $d = 1$ and $d = 3$ on the one hand, and $d = 2$ on the other hand, illustrates a general principle in d-dimensional Fourier analysis: a significant number of formulas that arise are simpler in the case of odd dimensions, compared to the corresponding situations in even dimensions. We will see several further examples of this below.

4 Radial symmetry and Bessel functions

We observed earlier that the Fourier transform of a radial function in $\mathbb{R}^d$ is also radial. In other words, if $f(x) = f_0(|x|)$ for some f_0, then

$\hat{f}(\xi) = F_0(|\xi|)$ for some F_0. A natural problem is to determine a relation between f_0 and F_0.

This problem has a simple answer in dimensions one and three. If $d = 1$ the relation we seek is

$$F_0(\rho) = 2\int_0^\infty \cos(2\pi\rho r) f_0(r)\, dr. \tag{9}$$

If we recall that $\mathbb{R}$ has only two rotations, the identity and multiplication by -1, we find that a function is radial precisely when it is even. Having made this observation it is easy to see that if f is radial, and $|\xi| = \rho$, then

$$\begin{aligned} F_0(\rho) = \hat{f}(|\xi|) &= \int_{-\infty}^\infty f(x) e^{-2\pi i x|\xi|}\, dx \\ &= \int_0^\infty f_0(r)(e^{-2\pi i r|\xi|} + e^{2\pi i r|\xi|})\, dr \\ &= 2\int_0^\infty \cos(2\pi\rho r) f_0(r)\, dr. \end{aligned}$$

In the case $d = 3$, the relation between f_0 and F_0 is also quite simple and given by the formula

$$F_0(\rho) = 2\rho^{-1}\int_0^\infty \sin(2\pi\rho r) f_0(r) r\, dr. \tag{10}$$

The proof of this identity is based on the formula for the Fourier transform of the surface element $d\sigma$ given in Lemma 3.5:

$$\begin{aligned} F_0(\rho) = \hat{f}(\xi) &= \int_{\mathbb{R}^3} f(x) e^{-2\pi i x\cdot\xi}\, dx \\ &= \int_0^\infty f_0(r)\int_{S^2} e^{-2\pi i r\gamma\cdot\xi} d\sigma(\gamma) r^2\, dr \\ &= \int_0^\infty f_0(r)\frac{2\sin(2\pi\rho r)}{\rho r} r^2\, dr \\ &= 2\rho^{-1}\int_0^\infty \sin(2\pi\rho r) f_0(r) r\, dr. \end{aligned}$$

More generally, the relation between f_0 and F_0 has a nice description in terms of a family of special functions that arise naturally in problems that exhibit radial symmetry.

The **Bessel function** of order $n \in \mathbb{Z}$, denoted $J_n(\rho)$, is defined as the nth Fourier coefficient of the function $e^{i\rho\sin\theta}$. So

$$J_n(\rho) = \frac{1}{2\pi}\int_0^{2\pi} e^{i\rho\sin\theta} e^{-in\theta}\, d\theta,$$

therefore

$$e^{i\rho\sin\theta} = \sum_{n=-\infty}^{\infty} J_n(\rho)e^{in\theta}.$$

As a result of this definition, we find that when $d = 2$, the relation between f_0 and F_0 is

$$F_0(\rho) = 2\pi \int_0^\infty J_0(2\pi r\rho) f_0(r) r\, dr. \tag{11}$$

Indeed, since $\hat{f}(\xi)$ is radial we take $\xi = (0, -\rho)$ so that

$$\begin{aligned}
\hat{f}(\xi) &= \int_{\mathbb{R}^2} f(x)e^{2\pi i x\cdot(0,\rho)}\, dx \\
&= \int_0^{2\pi}\int_0^\infty f_0(r)e^{2\pi i r\rho\sin\theta} r\, dr\, d\theta \\
&= 2\pi \int_0^\infty J_0(2\pi r\rho) f_0(r) r\, dr,
\end{aligned}$$

as desired.

In general, there are corresponding formulas relating f_0 and F_0 in $\mathbb{R}^d$ in terms of Bessel functions of order $d/2 - 1$ (see Problem 2). In even dimensions, these are the Bessel functions we have defined above. For odd dimensions, we need a more general definition of Bessel functions to encompass half-integral orders. Note that the formulas for the Fourier transform of radial functions give another illustration of the differences between odd and even dimensions. When $d = 1$ or $d = 3$ (as well as $d > 3$, d odd) the formulas are in terms of elementary functions, but this is not the case when d is even.

5 The Radon transform and some of its applications

Invented by Johann Radon in 1917, the integral transform we discuss next has many applications in mathematics and other sciences, including a significant achievement in medicine. To motivate the definitions and the central problem of reconstruction, we first present the close connection between the Radon transform and the development of X-ray scans (or CAT scans) in the theory of medical imaging. The solution of the reconstruction problem, and the introduction of new algorithms and faster computers, all contributed to a rapid development of computerized tomography. In practice, X-ray scans provide a “picture” of an internal organ, one that helps to detect and locate many types of abnormalities.

After a brief description of X-ray scans in two dimensions, we define the X-ray transform and formulate the basic problem of inverting this mapping. Although this problem has an explicit solution in $\mathbb{R}^2$, it is more complicated than the analogous problem in three dimensions, hence we give a complete solution of the reconstruction problem only in $\mathbb{R}^3$. Here we have another example where results are simpler in the odd-dimensional case than in the even-dimensional situation.

5.1 The X-ray transform in $\mathbb{R}^2$

Consider a two-dimensional object $\mathcal{O}$ lying in the plane $\mathbb{R}^2$, which we may think of as a planar cross section of a human organ.

First, we assume that $\mathcal{O}$ is homogeneous, and suppose that a very narrow beam of X-ray photons traverses this object.

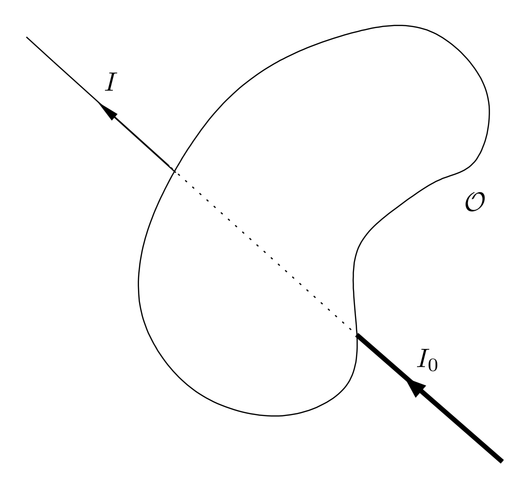


Figure 4. Attenuation of an X-ray beam

If I_0 and I denote the intensity of the beam before and after passing through $\mathcal{O}$, respectively, the following relation holds:

$$I = I_0 e^{-d\rho}.$$

Here d is the distance traveled by the beam in the object, and ρ denotes the **attenuation** coefficient (or absorption coefficient), which depends on the density and other physical characteristics of $\mathcal{O}$. If the object is not homogeneous, but consists of two materials with attenuation coefficients ρ_1 and ρ_2, then the observed decrease in the intensity of the beam is

given by

$$I = I_0 e^{-d_1\rho_1 - d_2\rho_2}$$

where d_1 and d_2 denote the distances traveled by the beam in each material. In the case of an arbitrary object whose density and physical characteristics vary from point to point, the attenuation factor is a function ρ in $\mathbb{R}^2$, and the above relations become

$$I = I_0 e^{\int_L \rho}.$$

Here L is the line in $\mathbb{R}^2$ traced by the beam, and $\int_L \rho$ denotes the line integral of ρ over L. Since we observe I and I_0, the data we gather after sending the X-ray beam through the object along the line L is the quantity

$$\int_L \rho.$$

Since we may initially send the beam in any given direction, we may calculate the above integral for every line in $\mathbb{R}^2$. We define the **X-ray transform** (or **Radon transform** in $\mathbb{R}^2$) of ρ by

$$X(\rho)(L) = \int_L \rho.$$

Note that this transform assigns to each appropriate function ρ on $\mathbb{R}^2$ (for example, $\rho \in \mathcal{S}(\mathbb{R}^2)$) another function $X(\rho)$ whose domain is the set of lines L in $\mathbb{R}^2$.

The unknown is ρ, and since our original interest lies precisely in the composition of the object, the problem now becomes to reconstruct the function ρ from the collected data, that is, its X-ray transform. We therefore pose the following *reconstruction* problem: Find a formula for ρ in terms of $X(\rho)$.

Mathematically, the problem asks for a formula giving the inverse of X. Does such an inverse even exist? As a first step, we pose the following simpler *uniqueness* question: If $X(\rho) = X(\rho')$, can we conclude that $\rho = \rho'$?

There is a reasonable *a priori* expectation that $X(\rho)$ actually determines ρ, as one can see by counting the dimensionality (or degrees of freedom) involved. A function ρ on $\mathbb{R}^2$ depends on two parameters (the x_1 and x_2 coordinates, for example). Similarly, the function $X(\rho)$, which is a function of lines L, is also determined by two parameters (for example, the slope of L and its x_2-intercept). In this sense, ρ and $X(\rho)$

convey an equivalent amount of information, so it is not unreasonable to suppose that $X(\rho)$ determines ρ.

While there is a satisfactory answer to the reconstruction problem, and a positive answer to the uniqueness question in $\mathbb{R}^2$, we shall forego giving them here. (However, see Exercise 13 and Problem 8.) Instead we shall deal with the analogous but simpler situation in $\mathbb{R}^3$.

Let us finally remark that in fact, one can sample the X-ray transform, and determine $X(\rho)(L)$ for only finitely many lines. Therefore, the reconstruction method implemented in practice is based not only on the general theory, but also on sampling procedures, numerical approximations, and computer algorithms. It turns out that a method used in developing effective relevant algorithms is the fast Fourier transform, which incidentally we take up in the next chapter.

5.2 The Radon transform in $\mathbb{R}^3$

The experiment described in the previous section applies in three dimensions as well. If $\mathcal{O}$ is an object in $\mathbb{R}^3$ determined by a function ρ which describes the density and physical characteristics of this object, sending an X-ray beam through $\mathcal{O}$ determines the quantity

$$\int_L \rho,$$

for every line in $\mathbb{R}^3$. In $\mathbb{R}^2$ this knowledge was enough to uniquely determine ρ, but in $\mathbb{R}^3$ we do not need as much data. In fact, by using the heuristic argument above of counting the number of degrees of freedom, we see that for functions ρ in $\mathbb{R}^3$ the number is three, while the number of parameters determining a line L in $\mathbb{R}^3$ is four (for example, two for the intercept in the (x_1, x_2) plane, and two more for the direction of the line). Thus in this sense, the problem is over-determined.

We turn instead to the natural mathematical generalization of the two-dimensional problem. Here we wish to determine the function in $\mathbb{R}^3$ by knowing its integral over all *planes*[3] in $\mathbb{R}^3$. To be precise, when we speak of a plane, we mean a plane not necessarily passing through the origin. If $\mathcal{P}$ is such a plane, we define the Radon transform $\mathcal{R}(f)$ by

$$\mathcal{R}(f)(\mathcal{P}) = \int_{\mathcal{P}} f.$$

To simplify our presentation, we shall follow our practice of assuming that we are dealing with functions in the class $\mathcal{S}(\mathbb{R}^3)$. However, many

[3]Note that the dimensionality associated with points on $\mathbb{R}^3$, and that for planes in $\mathbb{R}^3$, equals three in both cases.

of the results obtained below can be shown to be valid for much larger classes of functions.

First, we explain what we mean by the integral of f over a plane. The description we use for planes in $\mathbb{R}^3$ is the following: given a unit vector $\gamma \in S^2$ and a number $t \in \mathbb{R}$, we define the plane $\mathcal{P}_{t,\gamma}$ by

$$\mathcal{P}_{t,\gamma} = \{x \in \mathbb{R}^3 : x \cdot \gamma = t\}.$$

So we parametrize a plane by a unit vector γ orthogonal to it, and by its "distance" t to the origin (see Figure 5). Note that $\mathcal{P}_{t,\gamma} = \mathcal{P}_{-t,-\gamma}$, and we allow t to take negative values.

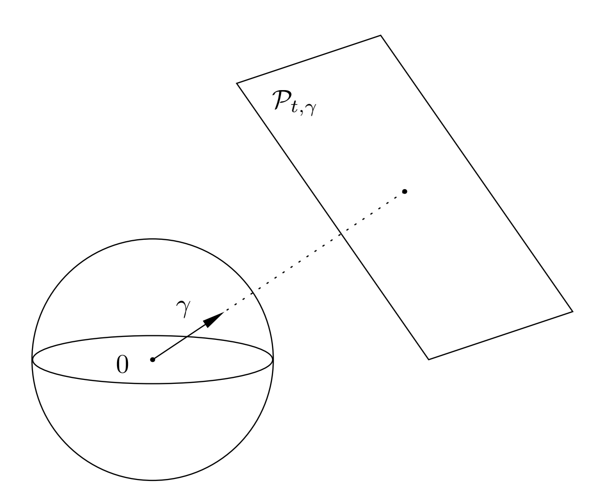


Figure 5. Description of a plane in $\mathbb{R}^3$

Given a function $f \in \mathcal{S}(\mathbb{R}^d)$, we need to make sense of its integral over $\mathcal{P}_{t,\gamma}$. We proceed as follows. Choose unit vectors e_1, e_2 so that e_1, e_2, γ is an orthonormal basis for $\mathbb{R}^3$. Then any $x \in \mathcal{P}_{t,\gamma}$ can be written uniquely as

$$x = t\gamma + u \quad \text{where} \quad u = u_1 e_1 + u_2 e_2 \quad \text{with } u_1, u_2 \in \mathbb{R}.$$

If $f \in \mathcal{S}(\mathbb{R}^3)$, we define

$$\int_{\mathcal{P}_{t,\gamma}} f = \int_{\mathbb{R}^2} f(t\gamma + u_1 e_1 + u_2 e_2)\, du_1\, du_2. \tag{12}$$

To be consistent, we must check that this definition is independent of the choice of the vectors e_1, e_2.

Proposition 5.1 *If $f \in \mathcal{S}(\mathbb{R}^3)$, then for each γ the definition of $\int_{\mathcal{P}_{t,\gamma}} f$ is independent of the choice of* e_1 *and* e_2*. Moreover*

$$\int_{-\infty}^{\infty} \left(\int_{\mathcal{P}_{t,\gamma}} f \right) dt = \int_{\mathbb{R}^3} f(x)\, dx.$$

Proof. If e_1', e_2' is another choice of basis vectors so that γ, e_1', e_2' is orthonormal, consider the rotation R in $\mathbb{R}^2$ which takes e_1 to e_1' and e_2 to e_2'. Changing variables $u' = R(u)$ in the integral proves that our definition (12) is independent of the choice of basis.

To prove the formula, let R denote the rotation which takes the standard basis of unit vectors[4] in $\mathbb{R}^3$ to γ, e_1, and e_2. Then

$$\begin{aligned} \int_{\mathbb{R}^3} f(x)\, dx &= \int_{\mathbb{R}^3} f(Rx)\, dx \\ &= \int_{\mathbb{R}^3} f(x_1\gamma + x_2 e_1 + x_3 e_2)\, dx_1\, dx_2\, dx_3 \\ &= \int_{-\infty}^{\infty} \left(\int_{\mathcal{P}_{t,\gamma}} f \right) dt. \end{aligned}$$

Remark. We digress to point out that the X-ray transform determines the Radon transform, since two-dimensional integrals can be expressed as iterated one-dimensional integrals. In other words, the knowledge of the integral of a function over all lines determines the integral of that function over any plane.

Having disposed of these preliminary matters, we turn to the study of the original problem. The **Radon transform** of a function $f \in \mathcal{S}(\mathbb{R}^3)$ is defined by

$$\mathcal{R}(f)(t, \gamma) = \int_{\mathcal{P}_{t,\gamma}} f.$$

In particular, we see that the Radon transform is a function on the set of planes in $\mathbb{R}^3$. From the parametrization given for a plane, we may equivalently think of $\mathcal{R}(f)$ as a function on the product $\mathbb{R} \times S^2 = \{(t, \gamma) : \ t \in \mathbb{R},\ \gamma \in S^2\}$, where S^2 denotes the unit sphere in $\mathbb{R}^3$. The relevant class of functions on $\mathbb{R} \times S^2$ consists of those that satisfy the Schwartz condition in t uniformly in γ. In other words, we define $\mathcal{S}(\mathbb{R} \times S^2)$ to be the space of all continuous functions $F(t, \gamma)$ that are indefinitely

[4]Here we are referring to the vectors $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.

differentiable in t, and that satisfy

$$\sup_{t\in\mathbb{R},\ \gamma\in S^2} |t|^k \left|\frac{d^\ell F}{\partial t^\ell}(t,\gamma)\right| < \infty \qquad \text{for all integers } k,\ell \geq 0.$$

Our goal is to solve the following problems.

Uniqueness problem: If $\mathcal{R}(f) = \mathcal{R}(g)$, then $f = g$.

Reconstruction problem: Express f in terms of $\mathcal{R}(f)$.

The solutions will be obtained by using the Fourier transform. In fact, the key point is a very elegant and essential relation between the Radon and Fourier transforms.

Lemma 5.2 *If $f \in \mathcal{S}(\mathbb{R}^3)$, then $\mathcal{R}(f)(t,\gamma) \in \mathcal{S}(\mathbb{R})$ for each fixed γ. Moreover,*

$$\widehat{\mathcal{R}}(f)(s,\gamma) = \hat{f}(s\gamma).$$

To be precise, $\hat{f}$ denotes the (three-dimensional) Fourier transform of f, while $\widehat{\mathcal{R}}(f)(s,\gamma)$ denotes the one-dimensional Fourier transform of $\mathcal{R}(f)(t,\gamma)$ as a function of t, with γ fixed.

Proof. Since $f \in \mathcal{S}(\mathbb{R}^3)$, for every positive integer N there is a constant $A_N < \infty$ so that

$$(1+|t|)^N(1+|u|)^N|f(t\gamma+u)| \leq A_N,$$

if we recall that $x = t\gamma + u$, where γ is orthogonal to u. Therefore, as soon as $N \geq 3$, we find

$$(1+|t|)^N\mathcal{R}(f)(t,\gamma) \leq A_N \int_{\mathbb{R}^2} \frac{du}{(1+|u|)^N} < \infty.$$

A similar argument for the derivatives shows that $\mathcal{R}(f)(t,\gamma) \in \mathcal{S}(\mathbb{R})$ for each fixed γ.

To establish the identity, we first note that

$$\begin{aligned}\widehat{\mathcal{R}}(f)(s,\gamma) &= \int_{-\infty}^{\infty}\left(\int_{\mathcal{P}_{t,\gamma}} f\right) e^{-2\pi ist}\,dt \\ &= \int_{-\infty}^{\infty}\int_{\mathbb{R}^2} f(t\gamma + u_1e_1 + u_2e_2)\,du_1\,du_2 e^{-2\pi ist}\,dt.\end{aligned}$$

However, since $\gamma \cdot u = 0$ and $|\gamma| = 1$, we may write

$$e^{-2\pi ist} = e^{-2\pi is\gamma\cdot(t\gamma+u)}.$$

As a result, we find that

$$\begin{aligned}\widehat{\mathcal{R}}(f)(s,\gamma) &= \int_{-\infty}^{\infty}\int_{\mathbb{R}^2} f(t\gamma + u_1e_1 + u_2e_2)e^{-2\pi is\gamma\cdot(t\gamma+u)}\,du_1\,du_2\,dt \\ &= \int_{-\infty}^{\infty}\int_{\mathbb{R}^2} f(t\gamma+u)e^{-2\pi is\gamma\cdot(t\gamma+u)}\,du\,dt.\end{aligned}$$

A final rotation from γ, e_1, e_2 to the standard basis in $\mathbb{R}^3$ proves that $\widehat{\mathcal{R}}(f)(s,\gamma) = \hat{f}(s\gamma)$, as desired.

As a consequence of this identity, we can answer the uniqueness question for the Radon transform in $\mathbb{R}^3$ in the affirmative.

Corollary 5.3 *If $f, g \in \mathcal{S}(\mathbb{R}^3)$ and $\mathcal{R}(f) = \mathcal{R}(g)$, then $f = g$.*

The proof of the corollary follows from an application of the lemma to the difference $f - g$ and use of the Fourier inversion theorem.

Our final task is to give the formula that allows us to recover f from its Radon transform. Since $\mathcal{R}(f)$ is a function on the set of planes in $\mathbb{R}^3$, and f is a function of the space variables $x \in \mathbb{R}^3$, to recover f we introduce the dual Radon transform, which passes from functions defined on planes to functions in $\mathbb{R}^3$.

Given a function F on $\mathbb{R} \times S^2$, we define its **dual Radon transform** by

$$\mathcal{R}^*(F)(x) = \int_{S^2} F(x\cdot\gamma, \gamma)\,d\sigma(\gamma). \tag{13}$$

Observe that a point x belongs to $\mathcal{P}_{t,\gamma}$ if and only if $x \cdot \gamma = t$, so the idea here is that given $x \in \mathbb{R}^3$, we obtain $\mathcal{R}^*(F)(x)$ by integrating F over the subset of all planes passing through x, that is,

$$\mathcal{R}^*(F)(x) = \int_{\{\mathcal{P}_{t,\gamma} \text{ such that } x\in\mathcal{P}_{t,\gamma}\}} F,$$

where the integral on the right is given the precise meaning in (13). We use the terminology "dual" because of the following observation. If $V_1 = \mathcal{S}(\mathbb{R}^3)$ with the usual Hermitian inner product

$$(f,g)_1 = \int_{\mathbb{R}^3} f(x)\overline{g(x)}\,dx,$$

and $V_2 = \mathcal{S}(\mathbb{R} \times S^2)$ with the Hermitian inner product

$$(F, G)_2 = \int_{\mathbb{R}} \int_{S^2} F(t, \gamma) \overline{G(t, \gamma)}\, d\sigma(\gamma)\, dt,$$

then

$$\mathcal{R} : V_1 \to V_2, \qquad \mathcal{R}^* : V_2 \to V_1,$$

with

$$(\mathcal{R}f, F)_2 = (f, \mathcal{R}^* F)_1. \tag{14}$$

The validity of this identity is not needed in the argument below, and its verification is left as an exercise for the reader.

We can now state the reconstruction theorem.

Theorem 5.4 *If $f \in \mathcal{S}(\mathbb{R}^3)$, then*

$$\triangle(\mathcal{R}^* \mathcal{R}(f)) = -8\pi^2 f.$$

We recall that $\triangle = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$ is the Laplacian.

Proof. By our previous lemma, we have

$$\mathcal{R}(f)(t, \gamma) = \int_{-\infty}^{\infty} \hat{f}(s\gamma) e^{2\pi i t s}\, ds.$$

Therefore

$$\mathcal{R}^* \mathcal{R}(f)(x) = \int_{S^2} \int_{-\infty}^{\infty} \hat{f}(s\gamma) e^{2\pi i x \cdot \gamma s}\, ds\, d\sigma(\gamma),$$

hence

$$\begin{aligned}
\triangle(\mathcal{R}^* \mathcal{R}(f))(x) &= \int_{S^2} \int_{-\infty}^{\infty} \hat{f}(s\gamma)(-4\pi^2 s^2) e^{2\pi i x \cdot \gamma s}\, ds\, d\sigma(\gamma) \\
&= -4\pi^2 \int_{S^2} \int_{-\infty}^{\infty} \hat{f}(s\gamma) e^{2\pi i x \cdot \gamma s} s^2\, ds\, d\sigma(\gamma) \\
&= -4\pi^2 \int_{S^2} \int_{-\infty}^{0} \hat{f}(s\gamma) e^{2\pi i x \cdot \gamma s} s^2\, ds\, d\sigma(\gamma) \\
&\qquad - 4\pi^2 \int_{S^2} \int_{0}^{\infty} \hat{f}(s\gamma) e^{2\pi i x \cdot \gamma s} s^2\, ds\, d\sigma(\gamma) \\
&= -8\pi^2 \int_{S^2} \int_{0}^{\infty} \hat{f}(s\gamma) e^{2\pi i x \cdot \gamma s} s^2\, ds\, d\sigma(\gamma) \\
&= -8\pi^2 f(x).
\end{aligned}$$

In the first line, we have differentiated under the integral sign and used the fact $\triangle(e^{2\pi ix\cdot\gamma s}) = (-4\pi^2 s^2)e^{2\pi ix\cdot\gamma s}$, since $|\gamma| = 1$. The last step follows from the formula for polar coordinates in $\mathbb{R}^3$ and the Fourier inversion theorem.

5.3 A note about plane waves

We conclude this chapter by briefly mentioning a nice connection between the Radon transform and solutions of the wave equation. This comes about in the following way. Recall that when $d = 1$, the solution of the wave equation can be expressed as the sum of traveling waves (see Chapter 1), and it is natural to ask if an analogue of such traveling waves exists in higher dimensions. The answer is as follows. Let F be a function of one variable, which we assume is sufficiently smooth (say C^2), and consider $u(x, t)$ defined by

$$u(x, t) = F((x \cdot \gamma) - t),$$

where $x \in \mathbb{R}^d$ and γ is a unit vector in $\mathbb{R}^d$. It is easy to verify directly that u is a solution of the wave equation in $\mathbb{R}^d$ (with $c = 1$). Such a solution is called a **plane wave**; indeed, notice that u is constant on every plane perpendicular to the direction γ, and as time t increases, the wave travels in the γ direction. (It should be remarked that plane waves are never functions in $\mathcal{S}(\mathbb{R}^d)$ when $d > 1$ because they are constant in directions perpendicular to γ).[5]

The basic fact is that when $d > 1$, the solution of the wave equation can be written as an integral (as opposed to sum, when $d = 1$) of plane waves; this can in fact be done via the Radon transform of the initial data f and g. For the relevant formulas when $d = 3$, see Problem 6.

6 Exercises

1. Suppose that R is a rotation in the plane $\mathbb{R}^2$, and let

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

denote its matrix with respect to the standard basis vectors $e_1 = (1, 0)$ and $e_2 = (0, 1)$.

(a) Write the conditions $R^t = R^{-1}$ and $\det(R) = \pm 1$ in terms of equations in a, b, c, d.

[5]Incidentally, this observation is further indication that a fuller treatment of the wave equation requires lifting the restriction that functions belong to $\mathcal{S}(\mathbb{R}^d)$.

(b) Show that there exists $\varphi \in \mathbb{R}$ such that $a + ib = e^{i\varphi}$.

(c) Conclude that if R is proper, then it can be expressed as $z \mapsto ze^{i\varphi}$, and if R is improper, then it takes the form $z \mapsto \overline{z}e^{i\varphi}$, where $\overline{z} = x - iy$.

2. Suppose that $R : \mathbb{R}^3 \to \mathbb{R}^3$ is a proper rotation.

(a) Show that $p(t) = \det(R - tI)$ is a polynomial of degree 3, and prove that there exists $\gamma \in S^2$ (where S^2 denotes the unit sphere in $\mathbb{R}^3$) with

$$R(\gamma) = \gamma.$$

[Hint: Use the fact that $p(0) > 0$ to see that there is $\lambda > 0$ with $p(\lambda) = 0$. Then $R - \lambda I$ is singular, so its kernel is non-trivial.]

(b) If $\mathcal{P}$ denotes the plane perpendicular to γ and passing through the origin, show that

$$R : \mathcal{P} \to \mathcal{P},$$

and that this linear map is a rotation.

3. Recall the formula

$$\int_{\mathbb{R}^d} F(x)\,dx = \int_{S^{d-1}} \int_0^\infty F(r\gamma) r^{d-1}\,dr\,d\sigma(\gamma).$$

Apply this to the special case when $F(x) = g(r)f(\gamma)$, where $x = r\gamma$, to prove that for any rotation R, one has

$$\int_{S^{d-1}} f(R(\gamma))\,d\sigma(\gamma) = \int_{S^{d-1}} f(\gamma)\,d\sigma(\gamma),$$

whenever f is a continuous function on the sphere S^{d-1}.

4. Let A_d and V_d denote the area and volume of the unit sphere and unit ball in $\mathbb{R}^d$, respectively.

(a) Prove the formula

$$A_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$$

so that $A_2 = 2\pi,\ A_3 = 4\pi,\ A_4 = 2\pi^2, \ldots$. Here $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\,dt$ is the Gamma function. [Hint: Use polar coordinates and the fact that $\int_{\mathbb{R}^d} e^{-\pi|x|^2}\,dx = 1$.]

(b) Show that $dV_d = A_d$, hence

$$V_d = \frac{\pi^{d/2}}{\Gamma(d/2+1)}.$$

In particular $V_2 = \pi$, $V_3 = 4\pi/3, \ldots$.

5. Let A be a $d \times d$ positive definite symmetric matrix with real coefficients. Show that

$$\int_{\mathbb{R}^d} e^{-\pi(x, A(x))}\, dx = (\det(A))^{-1/2}.$$

This generalizes the fact that $\int_{\mathbb{R}^d} e^{-\pi|x|^2}\, dx = 1$, which corresponds to the case where A is the identity.
[Hint: Apply the spectral theorem to write $A = RDR^{-1}$ where R is a rotation and, D is diagonal with entries $\lambda_1, \ldots, \lambda_d$, where $\{\lambda_i\}$ are the eigenvalues of A.]

6. Suppose $\psi \in \mathcal{S}(\mathbb{R}^d)$ satisfies $\int |\psi(x)|^2\, dx = 1$. Show that

$$\left(\int_{\mathbb{R}^d} |x|^2 |\psi(x)|^2\, dx\right)\left(\int_{\mathbb{R}^d} |\xi|^2 |\hat{\psi}(\xi)|^2\, d\xi\right) \geq \frac{d^2}{16\pi^2}.$$

This is the statement of the Heisenberg uncertainty principle in d dimensions.

7. Consider the time-dependent heat equation in $\mathbb{R}^d$:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_d^2}, \qquad \text{where } t > 0, \tag{15}$$

with boundary values $u(x,0) = f(x) \in \mathcal{S}(\mathbb{R}^d)$. If

$$\mathcal{H}_t^{(d)}(x) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t} = \int_{\mathbb{R}^d} e^{-4\pi^2 t|\xi|^2} e^{2\pi i x\cdot\xi}\, d\xi$$

is the d-dimensional **heat kernel**, show that the convolution

$$u(x,t) = (f * \mathcal{H}_t^{(d)})(x)$$

is indefinitely differentiable when $x \in \mathbb{R}^d$ and $t > 0$. Moreover, u solves (15), and is continuous up to the boundary $t = 0$ with $u(x,0) = f(x)$.

The reader may also wish to formulate the d-dimensional analogues of Theorem 2.1 and 2.3 in Chapter 5.

8. In Chapter 5, we found that a solution to the steady-state heat equation in the upper half-plane with boundary values f is given by the convolution $u = f * \mathcal{P}_y$ where the Poisson kernel is

$$\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2} \qquad \text{where } x \in \mathbb{R} \text{ and } y > 0.$$

More generally, one can calculate the d-dimensional Poisson kernel using the Fourier transform as follows.

(a) The **subordination principle** allows one to write expressions involving the function e^{-x} in terms of corresponding expressions involving the function e^{-x^2}. One form of this is the identity

$$e^{-\beta} = \int_0^\infty \frac{e^{-u}}{\sqrt{\pi u}} e^{-\beta^2/4u}\, du$$

when $\beta \geq 0$. Prove this identity with $\beta = 2\pi|x|$ by taking the Fourier transform of both sides.

(b) Consider the steady-state heat equation in the upper half-space $\{(x, y) : x \in \mathbb{R}^d,\ y > 0\}$

$$\sum_{j=1}^d \frac{\partial^2 u}{\partial x_j^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

with the Dirichlet boundary condition $u(x, 0) = f(x)$. A solution to this problem is given by the convolution $u(x, y) = (f * P_y^{(d)})(x)$ where $P_y^{(d)}(x)$ is the d-dimensional Poisson kernel

$$P_y^{(d)}(x) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi} e^{-2\pi|\xi| y}\, d\xi.$$

Compute $P_y^{(d)}(x)$ by using the subordination principle and the d-dimensional heat kernel. (See Exercise 7.) Show that

$$P_y^{(d)}(x) = \frac{\Gamma((d+1)/2)}{\pi^{(d+1)/2}} \frac{y}{(|x|^2 + y^2)^{(d+1)/2}}.$$

9. A **spherical wave** is a solution $u(x, t)$ of the Cauchy problem for the wave equation in $\mathbb{R}^d$, which as a function of x is radial. Prove that u is a spherical wave if and only if the initial data $f, g \in \mathcal{S}$ are both radial.

10. Let $u(x, t)$ be a solution of the wave equation, and let $E(t)$ denote the energy of this wave

$$E(t) = \int_{\mathbb{R}^d} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \sum_{j=1}^d \int_{\mathbb{R}^d} \left| \frac{\partial u}{\partial x_j}(x, t) \right|^2 dx.$$

We have seen that $E(t)$ is constant using Plancherel's formula. Give an alternate proof of this fact by differentiating the integral with respect to t and showing that

$$\frac{dE}{dt} = 0.$$

[Hint: Integrate by parts.]

11. Show that the solution of the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2}$$

subject to $u(x, 0) = f(x)$ and $\frac{\partial u}{\partial t}(x, 0) = g(x)$, where $f, g \in \mathcal{S}(\mathbb{R}^3)$, is given by

$$u(x,t) = \frac{1}{|S(x,t)|} \int_{S(x,t)} [tg(y) + f(y) + \nabla f(y) \cdot (y - x)] \, d\sigma(y),$$

where $S(x, t)$ denotes the sphere of center x and radius t, and $|S(x, t)|$ its area. This is an alternate expression for the solution of the wave equation given in Theorem 3.6. It is sometimes called **Kirchhoff's** formula.

12. Establish the identity (14) about the dual transform given in the text. In other words, prove that

$$\int_{\mathbb{R}} \int_{S^2} \mathcal{R}(f)(t,\gamma)\overline{F(t,\gamma)} d\sigma(\gamma) \, dt = \int_{\mathbb{R}^3} f(x)\overline{\mathcal{R}^*(F)(x)} \, dx \tag{16}$$

where $f \in \mathcal{S}(\mathbb{R}^3)$, $F \in \mathcal{S}(\mathbb{R} \times S^2)$, and

$$\mathcal{R}(f) = \int_{\mathcal{P}_{t,\gamma}} f \quad \text{and} \quad \mathcal{R}^*(F)(x) = \int_{S^2} F(x \cdot \gamma, \gamma) \, d\sigma(\gamma).$$

[Hint: Consider the integral

$$\int\int\int f(t\gamma + u_1 e_2 + u_2 e_2)\overline{F(t,\gamma)} \, dt \, d\sigma(\gamma) \, du_1 \, du_2.$$

Integrating first in u gives the left-hand side of (16), while integrating in u and t and setting $x = t\gamma + u_1 e_2 + u_2 e_2$ gives the right-hand side.]

13. For each (t, θ) with $t \in \mathbb{R}$ and $|\theta| \leq \pi$, let $L = L_{t,\theta}$ denote the line in the (x, y)-plane given by

$$x \cos\theta + y \sin\theta = t.$$

This is the line perpendicular to the direction $(\cos\theta, \sin\theta)$ at "distance" t from the origin (we allow negative t). For $f \in \mathcal{S}(\mathbb{R}^2)$ the X-ray transform or two-dimensional Radon transform of f is defined by

$$X(f)(t,\theta) = \int_{L_{t,\theta}} f = \int_{-\infty}^{\infty} f(t\cos\theta + u\sin\theta, t\sin\theta - u\cos\theta) \, du.$$

Calculate the X-ray transform of the function $f(x,y) = e^{-\pi(x^2+y^2)}$.

14. Let X be the X-ray transform. Show that if $f \in \mathcal{S}$ and $X(f) = 0$, then $f = 0$, by taking the Fourier transform in one variable.

15. For $F \in \mathcal{S}(\mathbb{R} \times S^1)$, define the **dual X-ray transform** $X^*(F)$ by integrating F over all lines that pass through the point (x,y) (that is, those lines $L_{t,\theta}$ with $x\cos\theta + y\sin\theta = t$):

$$X^*(F)(x,y) = \int F(x\cos\theta + y\sin\theta, \theta)\, d\theta.$$

Check that in this case, if $f \in \mathcal{S}(\mathbb{R}^2)$ and $F \in \mathcal{S}(\mathbb{R} \times S^1)$, then

$$\int\!\!\int X(f)(t,\theta)\overline{F(t,\theta)}\, dt\, d\theta = \int\!\!\int f(x,y)\overline{X^*(F)(x,y)}\, dx\, dy.$$

7 Problems

1. Let J_n denote the nth order Bessel function, for $n \in \mathbb{Z}$. Prove that

(a) $J_n(\rho)$ is real for all real ρ.

(b) $J_{-n}(\rho) = (-1)^n J_n(\rho)$.

(c) $2J_n'(\rho) = J_{n-1}(\rho) - J_{n+1}(\rho)$.

(d) $\left(\frac{2n}{\rho}\right) J_n(\rho) = J_{n-1}(\rho) + J_{n+1}(\rho)$.

(e) $(\rho^{-n}J_n(\rho))' = -\rho^{-n}J_{n+1}(\rho)$.

(f) $(\rho^n J_n(\rho))' = \rho^n J_{n-1}(\rho)$.

(g) $J_n(\rho)$ satisfies the second order differential equation

$$J_n''(\rho) + \rho^{-1}J_n'(\rho) + (1 - n^2/\rho^2)J_n(\rho) = 0.$$

(h) Show that

$$J_n(\rho) = \left(\frac{\rho}{2}\right)^n \sum_{m=0}^{\infty} (-1)^m \frac{\rho^{2m}}{2^{2m} m!(n+m)!}.$$

(i) Show that for all integers n and all real numbers a and b we have

$$J_n(a+b) = \sum_{\ell \in \mathbb{Z}} J_\ell(a) J_{n-\ell}(b).$$

2. Another formula for $J_n(\rho)$ that allows one to define Bessel functions for non-integral values of n, $(n > -1/2)$ is

$$J_n(\rho) = \frac{(\rho/2)^n}{\Gamma(n+1/2)\sqrt{\pi}} \int_{-1}^{1} e^{i\rho t}(1-t^2)^{n-(1/2)}\, dt.$$

(a) Check that the above formula agrees with the definition of $J_n(\rho)$ for integral $n \geq 0$. [Hint: Verify it for $n = 0$ and then check that both sides satisfy the recursion formula (e) in Problem 1.]

(b) Note that $J_{1/2}(\rho) = \sqrt{\frac{2}{\pi}}\rho^{-1/2}\sin\rho$.

(c) Prove that

$$\lim_{n\to -1/2} J_n(\rho) = \sqrt{\frac{2}{\pi}}\rho^{-1/2}\cos\rho.$$

(d) Observe that the formulas we have proved in the text giving F_0 in terms of f_0 (when describing the Fourier transform of a radial function) take the form

$$F_0(\rho) = 2\pi\rho^{-(d/2)+1}\int_0^{\infty} J_{(d/2)-1}(2\pi\rho r) f_0(r) r^{d/2}\, dr, \tag{17}$$

for $d = 1, 2$, and 3, if one uses the formulas above with the understanding that $J_{-1/2}(\rho) = \lim_{n\to -1/2} J_n(\rho)$. It turns out that the relation between F_0 and f_0 given by (17) is valid in all dimensions d.

3. We observed that the solution $u(x,t)$ of the Cauchy problem for the wave equation given by formula (3) depends only on the initial data on the base on the backward light cone. It is natural to ask if this property is shared by *any* solution of the wave equation. An affirmative answer would imply uniqueness of the solution.

Let $B(x_0, r_0)$ denote the closed ball in the hyperplane $t = 0$ centered at x_0 and of radius r_0. The **backward light cone** with base $B(x_0, r_0)$ is defined by

$$\mathcal{L}_{B(x_0,r_0)} = \{(x,t) \in \mathbb{R}^d \times \mathbb{R} : |x - x_0| \leq r_0 - t,\quad 0 \leq t \leq r_0\}.$$

Theorem *Suppose that $u(x,t)$ is a C^2 function on the closed upper half-plane $\{(x,t) :\ x \in \mathbb{R}^d, t \geq 0\}$ that solves the wave equation*

$$\frac{\partial^2 u}{\partial t^2} = \triangle u.$$

If $u(x,0) = \frac{\partial u}{\partial t}(x,0) = 0$ for all $x \in B(x_0,r_0)$, then $u(x,t) = 0$ for all $(x,t) \in \mathcal{L}_{B(x_0,r_0)}$.

In words, if the initial data of the Cauchy problem for the wave equation vanishes on a ball B, then *any* solution u of the problem vanishes in the backward light cone with base B. The following steps outline a proof of the theorem.

(a) Assume that u is real. For $0 \leq t \leq r_0$ let $B_t(x_0, r_0) = \{x : |x - x_0| \leq r_0 - t\}$, and also define

$$\nabla u(x,t) = \left(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_d}, \frac{\partial u}{\partial t}\right).$$

Now consider the energy integral

$$\begin{aligned} E(t) &= \frac{1}{2}\int_{B_t(x_0,r_0)} |\nabla u|^2 \, dx \\ &= \frac{1}{2}\int_{B_t(x_0,r_0)} \left(\frac{\partial u}{\partial t}\right)^2 + \sum_{j=1}^{d}\left(\frac{\partial u}{\partial x_j}\right)^2 dx. \end{aligned}$$

Observe that $E(t) \geq 0$ and $E(0) = 0$. Prove that

$$E'(t) = \int_{B_t(x_0,r_0)} \frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2} + \sum_{j=1}^{d} \frac{\partial u}{\partial x_j}\frac{\partial^2 u}{\partial x_j \partial t} \, dx - \frac{1}{2}\int_{\partial B_t(x_0,r_0)} |\nabla u|^2 \, d\sigma(\gamma).$$

(b) Show that

$$\frac{\partial}{\partial x_j}\left[\frac{\partial u}{\partial x_j}\frac{\partial u}{\partial t}\right] = \frac{\partial u}{\partial x_j}\frac{\partial^2 u}{\partial x_j \partial t} + \frac{\partial^2 u}{\partial x_j^2}\frac{\partial u}{\partial t}.$$

(c) Use the last identity, the divergence theorem, and the fact that u solves the wave equation to prove that

$$E'(t) = \int_{\partial B_t(x_0,r_0)} \sum_{j=1}^{d} \frac{\partial u}{\partial x_j}\frac{\partial u}{\partial t}\nu_j \, d\sigma(\gamma) - \frac{1}{2}\int_{\partial B_t(x_0,r_0)} |\nabla u|^2 \, d\sigma(\gamma),$$

where ν_j denotes the jth coordinate of the outward normal to $B_t(x_0, r_0)$.

(d) Use the Cauchy-Schwarz inequality to conclude that

$$\sum_{j=1}^{d} \frac{\partial u}{\partial x_j}\frac{\partial u}{\partial t}\nu_j \leq \frac{1}{2}|\nabla u|^2,$$

and as a result, $E'(t) \leq 0$. Deduce from this that $E(t) = 0$ and $u = 0$.

4.* There exist formulas for the solution of the Cauchy problem for the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_d^2} \qquad \text{with } u(x,0) = f(x) \text{ and } \frac{\partial u}{\partial t}(x,0) = g(x)$$

in $\mathbb{R}^d \times \mathbb{R}$ in terms of spherical means which generalize the formula given in the text for $d = 3$. In fact, the solution for even dimensions is deduced from that for odd dimensions, so we discuss this case first.

Suppose that $d > 1$ is odd and let $h \in \mathcal{S}(\mathbb{R}^d)$. The spherical mean of h on the ball centered at x of radius t is defined by

$$M_r h(x) = Mh(x,r) = \frac{1}{A_d} \int_{S^{d-1}} h(x - r\gamma)\, d\sigma(\gamma),$$

where A_d denotes the area of the unit sphere S^{d-1} in $\mathbb{R}^d$.

(a) Show that

$$\triangle_x Mh(x,r) = \left[\partial_r^2 + \frac{d-1}{r}\right] Mh(x,r),$$

where $\triangle_x$ denotes the Laplacian in the space variables x, and $\partial_r = \partial/\partial r$.

(b) Show that a twice differentiable function $u(x,t)$ satisfies the wave equation if and only if

$$\left[\partial_r^2 + \frac{d-1}{r}\right] Mu(x,r,t) = \partial_t^2 Mu(x,r,t),$$

where $Mu(x,r,t)$ denote the spherical means of the function $u(x,t)$.

(c) If $d = 2k+1$, define $T\varphi(r) = (r^{-1}\partial_r)^{k-1}[r^{2k-1}\varphi(r)]$, and let $\tilde{u} = TMu$. Then this function solves the one-dimensional wave equation for each fixed x:

$$\partial_t^2 \tilde{u}(x,r,t) = \partial_r^2 \tilde{u}(x,r,t).$$

One can then use d'Alembert's formula to find the solution $\tilde{u}(x,r,t)$ of this problem expressed in terms of the initial data.

(d) Now show that

$$u(x,t) = Mu(x,0,t) = \lim_{r \to 0} \frac{\tilde{u}(x,r,t)}{\alpha r}$$

where $\alpha = 1 \cdot 3 \cdots (d-2)$.

(e) Conclude that the solution of the Cauchy problem for the d-dimensional wave equation, when $d > 1$ is odd, is

$$u(x,t) = \frac{1}{1 \cdot 3 \cdots (d-2)} \left[\partial_t (t^{-1}\partial_t)^{(d-3)/2} \left(t^{d-2} M_t f(x)\right) + \right.$$

$$\left. (t^{-1}\partial_t)^{(d-3)/2} \left(t^{d-2} M_t g(x)\right)\right].$$

5.* The method of descent can be used to prove that the solution of the Cauchy problem for the wave equation in the case when d is even is given by the formula

$$u(x,t) = \frac{1}{1 \cdot 3 \cdots (d-2)} \left[\partial_t (t^{-1}\partial_t)^{(d-3)/2} \left(t^{d-2} \widetilde{M_t} f(x)\right) + \right.$$

$$\left. (t^{-1}\partial_t)^{(d-3)/2} \left(t^{d-2} \widetilde{M_t} g(x)\right)\right],$$

where $\widetilde{M_t}$ denotes the modified spherical means defined by

$$\widetilde{M_t} h(x) = \frac{2}{A_{d+1}} \int_{B^d} \frac{f(x+ty)}{\sqrt{1-|y|^2}}\, dy.$$

6.* Given initial data f and g of the form

$$f(x) = F(x \cdot \gamma) \quad \text{and} \quad g(x) = G(x \cdot \gamma),$$

check that the plane wave given by

$$u(x,t) = \frac{F(x\cdot\gamma + t) + F(x \cdot \gamma - t)}{2} + \frac{1}{2}\int_{x\cdot\gamma - t}^{x\cdot\gamma+t} G(s)\, ds$$

is a solution of the Cauchy problem for the d-dimensional wave equation.

In general, the solution is given as a superposition of plane waves. For the case $d = 3$, this can be expressed in terms of the Radon transform as follows. Let

$$\tilde{\mathcal{R}}(f)(t,\gamma) = -\frac{1}{8\pi^2}\left(\frac{d}{dt}\right)^2 R(f)(t,\gamma).$$

Then $u(x,t) =$

$$\frac{1}{2}\int_{S^2}\left[\tilde{\mathcal{R}}(f)(x\cdot\gamma - t,\gamma) + \tilde{\mathcal{R}}(f)(x\cdot\gamma + t,\gamma) + \int_{x\cdot\gamma-t}^{x\cdot\gamma+t} \tilde{\mathcal{R}}(g)(s,\gamma)\, ds\right] d\sigma(\gamma).$$

7. For every real number $a > 0$, define the operator $(-\triangle)^a$ by the formula

$$(-\triangle)^a f(x) = \int_{\mathbb{R}^d} (2\pi|\xi|)^{2a} \hat{f}(\xi) e^{2\pi i \xi \cdot x} \, d\xi$$

whenever $f \in \mathcal{S}(\mathbb{R}^d)$.

(a) Check that $(-\triangle)^a$ agrees with the usual definition of the ath power of $-\triangle$ (that is, a compositions of minus the Laplacian) when a is a positive integer.

(b) Verify that $(-\triangle)^a(f)$ is indefinitely differentiable.

(c) Prove that if a is not an integer, then in general $(-\triangle)^a(f)$ is not rapidly decreasing.

(d) Let $u(x, y)$ be the solution of the steady-state heat equation

$$\frac{\partial^2 u}{\partial y^2} + \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2} = 0, \qquad \text{with } u(x, 0) = f(x),$$

given by convolving f with the Poisson kernel (see Exercise 8). Check that

$$(-\triangle)^{1/2} f(x) = -\lim_{y \to 0} \frac{\partial u}{\partial y}(x, y),$$

and more generally that

$$(-\triangle)^{k/2} f(x) = (-1)^k \lim_{y \to 0} \frac{\partial^k u}{\partial y^k}(x, y)$$

for any positive integer k.

8.* The reconstruction formula for the Radon transform in $\mathbb{R}^d$ is as follows:

(a) When $d = 2$,

$$\frac{(-\triangle)^{1/2}}{4\pi} \mathcal{R}^*(\mathcal{R}(f)) = f,$$

where $(-\triangle)^{1/2}$ is defined in Problem 7.

(b) If the Radon transform and its dual are defined by analogy to the cases $d = 2$ and $d = 3$, then for general d,

$$\frac{(2\pi)^{1-d}}{2} (-\triangle)^{(d-1)/2} \mathcal{R}^*(\mathcal{R}(f)) = f.$$

7 Finite Fourier Analysis

> This past year has seen the birth, or rather the rebirth, of an exciting revolution in computing Fourier transforms. A class of algorithms known as the fast Fourier transform or FFT, is forcing a complete reassessment of many computational paths, not only in frequency analysis, but in any fields where problems can be reduced to Fourier transforms and/or convolutions...
>
> *C. Bingham and J. W. Tukey,* 1966

In the previous chapters we studied the Fourier series of functions on the circle and the Fourier transform of functions defined on the Euclidean space $\mathbb{R}^d$. The goal here is to introduce another version of Fourier analysis, now for functions defined on finite sets, and more precisely, on finite abelian groups. This theory is particularly elegant and simple since infinite sums and integrals are replaced by finite sums, and thus questions of convergence disappear.

In turning our attention to finite Fourier analysis, we begin with the simplest example, $\mathbb{Z}(N)$, where the underlying space is the (multiplicative) group of Nth roots of unity on the circle. This group can also be realized in additive form, as $\mathbb{Z}/N\mathbb{Z}$, the equivalence classes of integers modulo N. The group $\mathbb{Z}(N)$ arises as the natural approximation to the circle (as N tends to infinity) since in the first picture the points of $\mathbb{Z}(N)$ correspond to N points on the circle which are uniformly distributed. For this reason, in practical applications, the group $\mathbb{Z}(N)$ becomes a natural candidate for the storage of information of a function on the circle, and for the resulting numerical computations involving Fourier series. The situation is particularly nice when N is large and of the form $N = 2^n$. The computations of the Fourier coefficients now lead to the "fast Fourier transform," which exploits the fact that an induction in n requires only about $\log N$ steps to go from $N = 1$ to $N = 2^n$. This yields a substantial saving in time in practical applications.

In the second part of the chapter we undertake the more general theory of Fourier analysis on finite abelian groups. Here the fundamental example is the multiplicative group $\mathbb{Z}^*(q)$. The Fourier inversion formula

for $\mathbb{Z}^*(q)$ will be seen to be a key step in the proof of Dirichlet's theorem on primes in arithmetic progression, which we will take up in the next chapter.

1 Fourier analysis on $\mathbb{Z}(N)$

We turn to the group of Nth roots of unity. This group arises naturally as the simplest finite abelian group. It also gives a uniform partition of the circle, and is therefore a good choice if one wishes to sample appropriate functions on the circle. Moreover, this partition gets finer as N tends to infinity, and one might expect that the discrete Fourier theory that we discuss here tends to the continuous theory of Fourier series on the circle. In a broad sense, this is the case, although this aspect of the problem is not one that we develop.

1.1 The group $\mathbb{Z}(N)$

Let N be a positive integer. A complex number z is an Nth **root of unity** if $z^N = 1$. The set of Nth roots of unity is precisely

$$\left\{1, e^{2\pi i/N}, e^{2\pi i2/N}, \ldots, e^{2\pi i(N-1)/N}\right\}.$$

Indeed, suppose that $z^N = 1$ with $z = re^{i\theta}$. Then we must have $r^N e^{iN\theta} = 1$, and taking absolute values yields $r = 1$. Therefore $e^{iN\theta} = 1$, and this means that $N\theta = 2\pi k$ where $k \in \mathbb{Z}$. So if $\zeta = e^{2\pi i/N}$ we find that ζ^k exhausts all the Nth roots of unity. However, notice that $\zeta^N = 1$ so if n and m differ by an integer multiple of N, then $\zeta^n = \zeta^m$. In fact, it is clear that

$$\zeta^n = \zeta^m \quad \text{if and only if} \quad n - m \text{ is divisible by } N.$$

We denote the set of all N^{th} roots of unity by $\mathbb{Z}(N)$. The fact that this set gives a uniform partition of the circle is clear from its definition. Note that the set $\mathbb{Z}(N)$ satisfies the following properties:

(i) If $z, w \in \mathbb{Z}(N)$, then $zw \in \mathbb{Z}(N)$ and $zw = wz$.

(ii) $1 \in \mathbb{Z}(N)$.

(iii) If $z \in \mathbb{Z}(N)$, then $z^{-1} = 1/z \in \mathbb{Z}(N)$ and of course $zz^{-1} = 1$.

As a result we can conclude that $\mathbb{Z}(N)$ is an abelian group under complex multiplication. The appropriate definitions are set out in detail later in Section 2.1.

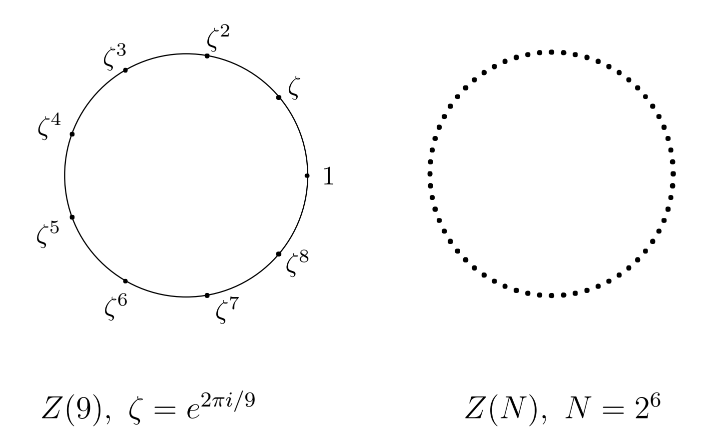


Figure 1. The group of Nth roots of unity when $N = 9$ and $N = 2^6 = 64$

There is another way to visualize the group $\mathbb{Z}(N)$. This consists of choosing the integer power of ζ that determines each root of unity. We observed above that this integer is not unique since $\zeta^n = \zeta^m$ whenever n and m differ by an integer multiple of N. Naturally, we might select the integer which satisfies $0 \leq n \leq N-1$. Although this choice is perfectly reasonable in terms of "sets," we ask what happens when we multiply roots of unity. Clearly, we must add the corresponding integers since $\zeta^n\zeta^m = \zeta^{n+m}$ but nothing guarantees that $0 \leq n+m \leq N-1$. In fact, if $\zeta^n\zeta^m = \zeta^k$ with $0 \leq k \leq N-1$, then $n+m$ and k differ by an integer multiple of N. So, to find the integer in $[0, N-1]$ corresponding to the root of unity $\zeta^n\zeta^m$, we see that after adding the integers n and m we must reduce modulo N, that is, find the unique integer $0 \leq k \leq N-1$ so that $(n+m)-k$ is an integer multiple of N.

An equivalent approach is to associate to each root of unity ω the class of integers n so that $\zeta^n = \omega$. Doing so for each root of unity we obtain a partition of the integers in N disjoint infinite classes. To add two of these classes, choose any integer in each one of them, say n and m, respectively, and define the sum of the classes to be the class which contains the integer $n+m$.

We formalize the above notions. Two integers x and y are **congruent modulo** N if the difference $x-y$ is divisible by N, and we write $x \equiv y \mod N$. In other words, this means that x and y differ by an integer multiple of N. It is an easy exercise to check the following three properties:

- $x \equiv x \mod N$ for all integers x.
- If $x \equiv y \mod N$, then $y \equiv x \mod N$.
- If $x \equiv y \mod N$ and $y \equiv z \mod N$, then $x \equiv z \mod N$.

The above defines an equivalence relation on $\mathbb{Z}$. Let $R(x)$ denote the equivalence class, or residue class, of the integer x. Any integer of the form $x + kN$ with $k \in \mathbb{Z}$ is an element (or "representative") of $R(x)$. In fact, there are precisely N equivalence classes, and each class has a unique representative between 0 and $N - 1$. We may now add equivalence classes by defining

$$R(x) + R(y) = R(x + y).$$

This definition is of course independent of the representatives x and y because if $x' \in R(x)$ and $y' \in R(y)$, then one checks easily that $x' + y' \in R(x + y)$. This turns the set of equivalence classes into an abelian group called the **group of integers modulo** N, which is sometimes denoted by $\mathbb{Z}/N\mathbb{Z}$. The association

$$R(k) \longleftrightarrow e^{2\pi ik/N}$$

gives a correspondence between the two abelian groups, $\mathbb{Z}/N\mathbb{Z}$ and $\mathbb{Z}(N)$. Since the operations are respected, in the sense that addition of integers modulo N becomes multiplication of complex numbers, we shall also denote the group of integers modulo N by $\mathbb{Z}(N)$. Observe that $0 \in \mathbb{Z}/N\mathbb{Z}$ corresponds to 1 on the unit circle.

Let V and W denote the vector spaces of complex-valued functions on the group of integers modulo N and the Nth roots of unity, respectively. Then, the identification given above carries over to V and W as follows:

$$F(k) \longleftrightarrow f(e^{2\pi ik/N}),$$

where F is a function on the integers modulo N and f is a function on the Nth roots of unity.

From now on, we write $\mathbb{Z}(N)$ but think of either the group of integers modulo N or the group of Nth roots of unity.

1.2 Fourier inversion theorem and Plancherel identity on $\mathbb{Z}(N)$

The first and most crucial step in developing Fourier analysis on $\mathbb{Z}(N)$ is to find the functions which correspond to the exponentials $e_n(x) = e^{2\pi inx}$ in the case of the circle. Some important properties of these exponentials are:

(i) $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal set for the inner product (1) (in Chapter 3) on the space of Riemann integrable functions on the circle.

(ii) Finite linear combinations of the e_n's (the trigonometric polynomials) are dense in the space of continuous functions on the circle.

(iii) $e_n(x+y) = e_n(x)e_n(y)$.

On $\mathbb{Z}(N)$, the appropriate analogues are the N functions $e_0, \ldots, e_{N-1}$ defined by

$$e_\ell(k) = \zeta^{\ell k} = e^{2\pi i \ell k/N} \qquad \text{for } \ell = 0, \ldots, N-1 \text{ and } k = 0, \ldots, N-1,$$

where $\zeta = e^{2\pi i/N}$. To understand the parallel with (i) and (ii), we can think of the complex-valued functions on $\mathbb{Z}(N)$ as a vector space V, endowed with the Hermitian inner product

$$(F, G) = \sum_{k=0}^{N-1} F(k)\overline{G(k)}$$

and associated norm

$$\|F\|^2 = \sum_{k=0}^{N-1} |F(k)|^2.$$

Lemma 1.1 *The family $\{e_0, \ldots, e_{N-1}\}$ is orthogonal. In fact,*

$$(e_m, e_\ell) = \begin{cases} N & \text{if } m = \ell, \\ 0 & \text{if } m \neq \ell. \end{cases}$$

Proof. We have

$$(e_m, e_\ell) = \sum_{k=0}^{N-1} \zeta^{mk}\zeta^{-\ell k} = \sum_{k=0}^{N-1} \zeta^{(m-\ell)k}.$$

If $m = \ell$, each term in the sum is equal to 1, and the sum equals N. If $m \neq \ell$, then $q = \zeta^{m-\ell}$ is not equal to 1, and the usual formula

$$1 + q + q^2 + \cdots + q^{N-1} = \frac{1-q^N}{1-q}$$

shows that $(e_m, e_\ell) = 0$, because $q^N = 1$.

Since the N functions $e_0, \ldots, e_{N-1}$ are orthogonal, they must be linearly independent, and since the vector space V is N-dimensional, we

conclude that $\{e_0, \ldots, e_{N-1}\}$ is an orthogonal basis for V. Clearly, property (iii) also holds, that is, $e_\ell(k+m) = e_\ell(k)e_\ell(m)$ for all ℓ, and all $k, m \in \mathbb{Z}(N)$.

By the lemma each vector e_ℓ has norm $\sqrt{N}$, so if we define

$$e_\ell^* = \frac{1}{\sqrt{N}}\, e_\ell,$$

then $\{e_0^*, \ldots, e_{N-1}^*\}$ is an orthonormal basis for V. Hence for any $F \in V$ we have

$$(1) \qquad F = \sum_{n=0}^{N-1} (F, e_n^*)e_n^* \quad \text{as well as} \quad \|F\|^2 = \sum_{n=0}^{N-1} |(F, e_n^*)|^2.$$

If we define the nth **Fourier coefficient** of F by

$$a_n = \frac{1}{N} \sum_{k=0}^{N-1} F(k) e^{-2\pi i k n/N},$$

the above observations give the following fundamental theorem which is the $\mathbb{Z}(N)$ version of the Fourier inversion and the Parseval-Plancherel formulas.

Theorem 1.2 *If F is a function on $\mathbb{Z}(N)$, then*

$$F(k) = \sum_{n=0}^{N-1} a_n e^{2\pi i n k/N}.$$

Moreover,

$$\sum_{n=0}^{N-1} |a_n|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |F(k)|^2.$$

The proof follows directly from (1) once we observe that

$$a_n = \frac{1}{N}\,(F, e_n) = \frac{1}{\sqrt{N}}\,(F, e_n^*).$$

Remark. It is possible to recover the Fourier inversion on the circle for sufficiently smooth functions (say C^2) by letting $N \to \infty$ in the finite model $\mathbb{Z}(N)$ (see Exercise 3).

1.3 The fast Fourier transform

The fast Fourier transform is a method that was developed as a means of calculating efficiently the Fourier coefficients of a function F on $\mathbb{Z}(N)$.

The problem, which arises naturally in numerical analysis, is to determine an algorithm that minimizes the amount of time it takes a computer to calculate the Fourier coefficients of a given function on $\mathbb{Z}(N)$. Since this amount of time is roughly proportional to the number of operations the computer must perform, our problem becomes that of minimizing the number of operations necessary to obtain all the Fourier coefficients $\{a_n\}$ given the values of F on $\mathbb{Z}(N)$. By operations we mean either an addition or a multiplication of complex numbers.

We begin with a naive approach to the problem. Fix N, and suppose that we are given $F(0), \ldots, F(N-1)$ and $\omega_N = e^{-2\pi i/N}$. If we denote by $a_k^N(F)$ the kth Fourier coefficient of F on $\mathbb{Z}(N)$, then by definition

$$a_k^N(F) = \frac{1}{N}\sum_{r=0}^{N-1} F(r)\omega_N^{kr},$$

and crude estimates show that the number of operations needed to calculate all Fourier coefficients is $\leq 2N^2 + N$. Indeed, it takes at most $N-2$ multiplications to determine $\omega_N^2, \ldots, \omega_N^{N-1}$, and each coefficient a_k^N requires $N+1$ multiplications and $N-1$ additions.

We now present the **fast Fourier transform**, an algorithm that improves the bound $O(N^2)$ obtained above. Such an improvement is possible if, for example, we restrict ourselves to the case where the partition of the circle is dyadic, that is, $N = 2^n$. (See also Exercise 9.)

Theorem 1.3 *Given $\omega_N = e^{-2\pi i/N}$ with $N = 2^n$, it is possible to calculate the Fourier coefficients of a function on $\mathbb{Z}(N)$ with at most*

$$4 \cdot 2^n n = 4N \log_2(N) = O(N \log N)$$

operations.

The proof of the theorem consists of using the calculations for M division points, to obtain the Fourier coefficients for $2M$ division points. Since we choose $N = 2^n$, we obtain the desired formula as a consequence of a recurrence which involves $n = O(\log N)$ steps.

Let $\#(M)$ denote the minimum number of operations needed to calculate all the Fourier coefficients of any function on $\mathbb{Z}(M)$. The key to the proof of the theorem is contained in the following recursion step.

Lemma 1.4 *If we are given* $\omega_{2M} = e^{-2\pi i/(2M)}$*, then*

$$\#(2M) \leq 2\#(M) + 8M.$$

Proof. The calculation of $\omega_{2M}, \ldots, \omega_{2M}^{2M}$ requires no more than $2M$ operations. Note that in particular we get $\omega_M = e^{-2\pi i/M} = \omega_{2M}^2$. The main idea is that for any given function F on $\mathbb{Z}(2M)$, we consider two functions F_0 and F_1 on $\mathbb{Z}(M)$ defined by

$$F_0(r) = F(2r) \quad \text{and} \quad F_1(r) = F(2r+1).$$

We assume that it is possible to calculate the Fourier coefficients of F_0 and F_1 in no more than $\#(M)$ operations each. If we denote the Fourier coefficients corresponding to the groups $\mathbb{Z}(2M)$ and $\mathbb{Z}(M)$ by a_k^{2M} and a_k^M, respectively, then we have

$$a_k^{2M}(F) = \frac{1}{2}\left(a_k^M(F_0) + a_k^M(F_1)\omega_{2M}^k\right).$$

To prove this, we sum over odd and even integers in the definition of the Fourier coefficient $a_k^{2M}(F)$, and find

$$\begin{aligned}
a_k^{2M}(F) &= \frac{1}{2M}\sum_{r=0}^{2M-1} F(r)\omega_{2M}^{kr} \\
&= \frac{1}{2}\left(\frac{1}{M}\sum_{\ell=0}^{M-1} F(2\ell)\omega_{2M}^{k(2\ell)} + \frac{1}{M}\sum_{m=0}^{M-1} F(2m+1)\omega_{2M}^{k(2m+1)}\right) \\
&= \frac{1}{2}\left(\frac{1}{M}\sum_{\ell=0}^{M-1} F_0(\ell)\omega_M^{k\ell} + \frac{1}{M}\sum_{m=0}^{M-1} F_1(m)\omega_M^{km}\omega_{2M}^k\right),
\end{aligned}$$

which establishes our assertion.

As a result, knowing $a_k^M(F_0)$, $a_k^M(F_1)$, and ω_{2M}^k, we see that each $a_k^{2M}(F)$ can be computed using no more than three operations (one addition and two multiplications). So

$$\#(2M) \leq 2M + 2\#(M) + 3 \times 2M = 2\#(M) + 8M,$$

and the proof of the lemma is complete.

An induction on n, where $N = 2^n$, will conclude the proof of the theorem. The initial step $n = 1$ is easy, since $N = 2$ and the two Fourier coefficients are

$$a_0^N(F) = \frac{1}{2}\left(F(1) + F(-1)\right) \quad \text{and} \quad a_1^N(F) = \frac{1}{2}\left(F(1) + (-1)F(-1)\right).$$

Calculating these Fourier coefficients requires no more than five operations, which is less than $4 \times 2 = 8$. Suppose the theorem is true up to $N = 2^{n-1}$ so that $\#(N) \leq 4 \cdot 2^{n-1}(n-1)$. By the lemma we must have

$$\#(2N) \leq 2 \cdot 4 \cdot 2^{n-1}(n-1) + 8 \cdot 2^{n-1} = 4 \cdot 2^n n,$$

which concludes the inductive step and the proof of the theorem.

2 Fourier analysis on finite abelian groups

The main goal in the rest of this chapter is to generalize the results about Fourier series expansions obtained in the special case of $\mathbb{Z}(N)$.

After a brief introduction to some notions related to finite abelian groups, we turn to the important concept of a character. In our setting, we find that characters play the same role as the exponentials $e_0, \ldots, e_{N-1}$ on the group $\mathbb{Z}(N)$, and thus provide the key ingredient in the development of the theory on arbitrary finite abelian groups. In fact, it suffices to prove that a finite abelian group has "enough" characters, and this leads automatically to the desired Fourier theory.

2.1 Abelian groups

An **abelian group** (or commutative group) is a set G together with a binary operation on pairs of elements of G, $(a, b) \mapsto a \cdot b$, that satisfies the following conditions:

(i) *Associativity*: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in G$.

(ii) *Identity*: There exists an element $u \in G$ (often written as either 1 or 0) such that $a \cdot u = u \cdot a = a$ for all $a \in G$.

(iii) *Inverses*: For every $a \in G$, there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = u$.

(iv) *Commutativity*: For $a, b \in G$, we have $a \cdot b = b \cdot a$.

We leave as simple verifications the facts that the identity element and inverses are unique.

Warning. In the definition of an abelian group, we used the "multiplicative" notation for the operation in G. Sometimes, one uses the "additive" notation $a + b$ and $-a$, instead of $a \cdot b$ and a^{-1}. There are times when one notation may be more appropriate than the other, and the examples below illustrate this point. The same group may have different interpretations, one where the multiplicative notation is more suggestive, and another where it is natural to view the group with addition, as the operation.

Examples of abelian groups

- The set of real numbers $\mathbb{R}$ with the usual addition. The identity is 0 and the inverse of x is $-x$.

  Also, $\mathbb{R} - \{0\}$ and $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$ equipped, with the standard multiplication, are abelian groups. In both cases the unit is 1 and the inverse of x is $1/x$.

- With the usual addition, the set of integers $\mathbb{Z}$ is an abelian group. However, $\mathbb{Z} - \{0\}$ is not an abelian group with the standard multiplication, since, for example, 2 does not have a multiplicative inverse in $\mathbb{Z}$. In contrast, $\mathbb{Q} - \{0\}$ is an abelian group with the standard multiplication.

- The unit circle S^1 in the complex plane. If we view the circle as the set of points $\{e^{i\theta} : \theta \in \mathbb{R}\}$, the group operation is the standard multiplication of complex numbers. However, if we identify points on S^1 with their angle θ, then S^1 becomes $\mathbb{R}$ modulo 2π, where the operation is addition modulo 2π.

- $\mathbb{Z}(N)$ is an abelian group. Viewed as the Nth roots of unity on the circle, $\mathbb{Z}(N)$ is a group under multiplication of complex numbers. However, if $\mathbb{Z}(N)$ is interpreted as $\mathbb{Z}/N\mathbb{Z}$, the integers modulo N, then it is an abelian group where the operation is addition modulo N.

- The last example consists of $\mathbb{Z}^*(q)$. This group is defined as the set of all integers modulo q that have *multiplicative* inverses, with the group operation being multiplication modulo q. This important example is discussed in more detail below.

A **homomorphism** between two abelian groups G and H is a map $f : G \to H$ which satisfies the property

$$f(a \cdot b) = f(a) \cdot f(b),$$

where the dot on the left-hand side is the operation in G, and the dot on the right-hand side the operation in H.

We say that two groups G and H are **isomorphic**, and write $G \approx H$, if there is a bijective homomorphism from G to H. Equivalently, G and H are isomorphic if there exists another homomorphism $\tilde{f} : H \to G$, so that for all $a \in G$ and $b \in H$

$$(\tilde{f} \circ f)(a) = a \quad \text{and} \quad (f \circ \tilde{f})(b) = b.$$

Roughly speaking, isomorphic groups describe the "same" object because they have the same underlying group structure (which is really all that matters); however, their particular notational representations might be different.

EXAMPLE 1. A pair of isomorphic abelian groups arose already when we considered the group $\mathbb{Z}(N)$. In one representation it was given as the multiplicative group of Nth roots of unity in $\mathbb{C}$. In a second representation it was the additive group $\mathbb{Z}/N\mathbb{Z}$ of residue classes of integers modulo N. The mapping $n \mapsto R(n)$, which associates to a root of unity $z = e^{2\pi in/N} = \zeta^n$ the residue class in $\mathbb{Z}/N\mathbb{Z}$ determined by n, provides an isomorphism between the two different representations.

EXAMPLE 2. In parallel with the previous example, we see that the circle (with multiplication) is isomorphic to the real numbers modulo 2π (with addition).

EXAMPLE 3. The properties of the exponential and logarithm guarantee that

$$\exp : \mathbb{R} \to \mathbb{R}^+ \quad \text{and} \quad \log : \mathbb{R}^+ \to \mathbb{R}$$

are two homomorphisms that are inverses of each other. Thus $\mathbb{R}$ (with addition) and $\mathbb{R}^+$ (with multiplication) are isomorphic.

In what follows, we are primarily interested in abelian groups that are finite. In this case, we denote by $|G|$ the number of elements in G, and call $|G|$ the **order** of the group. For example, the order of $\mathbb{Z}(N)$ is N.

A few additional remarks are in order:

- If G_1 and G_2 are two finite abelian groups, their **direct product** $G_1 \times G_2$ is the group whose elements are pairs (g_1, g_2) with $g_1 \in G_1$ and $g_2 \in G_2$. The operation in $G_1 \times G_2$ is then defined by

$$(g_1, g_2) \cdot (g_1', g_2') = (g_1 \cdot g_1', g_2 \cdot g_2').$$

  Clearly, if G_1 and G_2 are finite abelian groups, then so is $G_1 \times G_2$. The definition of direct product generalizes immediately to the case of finitely many factors $G_1 \times G_2 \times \cdots \times G_n$.

- The structure theorem for finite abelian groups states that such a group is isomorphic to a direct product of groups of the type $\mathbb{Z}(N)$; see Problem 2. This is a nice result which gives us an overview of the class of all finite abelian groups. However, since we shall not use this theorem below, we omit its proof.

We now discuss briefly the examples of abelian groups that play a central role in the proof of Dirichlet's theorem in the next chapter.

The group $\mathbb{Z}^*(q)$

Let q be a positive integer. We see that multiplication in $\mathbb{Z}(q)$ can be unambiguously defined, because if n is congruent to n' and m is congruent to m' (both modulo q), then nm is congruent to $n'm'$ modulo q. An integer $n \in \mathbb{Z}(q)$ is a **unit** if there exists an integer $m \in \mathbb{Z}(q)$ so that

$$nm \equiv 1 \mod q.$$

The set of all units in $\mathbb{Z}(q)$ is denoted by $\mathbb{Z}^*(q)$, and it is clear from our definition that $\mathbb{Z}^*(q)$ is an abelian group under *multiplication* modulo q. Thus within the additive group $\mathbb{Z}(q)$ lies a set $\mathbb{Z}^*(q)$ that is a group under multiplication. An alternative characterization of $\mathbb{Z}^*(q)$ will be given in the next chapter, as those elements in $\mathbb{Z}(q)$ that are relatively prime to q.

EXAMPLE 4. The group of units in $\mathbb{Z}(4) = \{0, 1, 2, 3\}$ is

$$\mathbb{Z}^*(4) = \{1, 3\}.$$

This reflects the fact that odd integers are divided into two classes depending on whether they are of the form $4k + 1$ or $4k + 3$. In fact, $\mathbb{Z}^*(4)$ is isomorphic to $\mathbb{Z}(2)$. Indeed, we can make the following association:

$$\begin{array}{ccc} \mathbb{Z}^*(4) & & \mathbb{Z}(2) \\ 1 & \longleftrightarrow & 0 \\ 3 & \longleftrightarrow & 1 \end{array}$$

and then notice that multiplication in $\mathbb{Z}^*(4)$ corresponds to addition in $\mathbb{Z}(2)$.

EXAMPLE 5. The units in $\mathbb{Z}(5)$ are

$$\mathbb{Z}^*(5) = \{1, 2, 3, 4\}.$$

Moreover, $\mathbb{Z}^*(5)$ is isomorphic to $\mathbb{Z}(4)$ with the following identification:

$$\begin{array}{ccc} \mathbb{Z}^*(5) & & \mathbb{Z}(4) \\ 1 & \longleftrightarrow & 0 \\ 2 & \longleftrightarrow & 1 \\ 3 & \longleftrightarrow & 3 \\ 4 & \longleftrightarrow & 2 \end{array}$$

EXAMPLE 6. The units in $\mathbb{Z}(8) = \{0, 1, 2, 3, 4, 5, 6, 7\}$ are given by

$$\mathbb{Z}^*(8) = \{1, 3, 5, 7\}.$$

In fact, $\mathbb{Z}^*(8)$ is isomorphic to the direct product $\mathbb{Z}(2) \times \mathbb{Z}(2)$. In this case, an isomorphism between the groups is given by the identification

$\mathbb{Z}^*(8)$		$\mathbb{Z}(2) \times \mathbb{Z}(2)$
1	$\longleftrightarrow$	$(0, 0)$
3	$\longleftrightarrow$	$(1, 0)$
5	$\longleftrightarrow$	$(0, 1)$
7	$\longleftrightarrow$	$(1, 1)$

2.2 Characters

Let G be a finite abelian group (with the multiplicative notation) and S^1 the unit circle in the complex plane. A **character** on G is a complex-valued function $e : G \to S^1$ which satisfies the following condition:

$$e(a \cdot b) = e(a)e(b) \quad \text{for all } a, b \in G. \tag{2}$$

In other words, a character is a homomorphism from G to the circle group. The **trivial** or **unit character** is defined by $e(a) = 1$ for all $a \in G$.

Characters play an important role in the context of finite Fourier series, primarily because the multiplicative property (2) generalizes the analogous identity for the exponential functions on the circle and the law

$$e_\ell(k + m) = e_\ell(k)e_\ell(m),$$

which held for the exponentials $e_0, \ldots, e_{N-1}$ used in the Fourier theory on $\mathbb{Z}(N)$. There we had $e_\ell(k) = \zeta^{\ell k} = e^{2\pi i \ell k/N}$, with $0 \le \ell \le N - 1$ and $k \in \mathbb{Z}(N)$, and in fact, the functions $e_0, \ldots, e_{N-1}$ are precisely all the characters of the group $\mathbb{Z}(N)$.

If G is a finite abelian group, we denote by $\hat{G}$ the set of all characters of G, and observe next that this set inherits the structure of an abelian group.

Lemma 2.1 *The set $\hat{G}$ is an abelian group under multiplication defined by*

$$(e_1 \cdot e_2)(a) = e_1(a)e_2(a) \quad \textit{for all } a \in G.$$

The proof of this assertion is straightforward if one observes that the trivial character plays the role of the unit. We call $\hat{G}$ the **dual group** of G.

In light of the above analogy between characters for a general abelian group and the exponentials on $\mathbb{Z}(N)$, we gather several more examples of groups and their duals. This provides further evidence of the central role played by characters. (See Exercises 4, 5, and 6.)

Example 1. If $G = \mathbb{Z}(N)$, all characters of G take the form $e_\ell(k) = \zeta^{\ell k} = e^{2\pi i \ell k/N}$ for some $0 \leq \ell \leq N-1$, and it is easy to check that $e_\ell \mapsto \ell$ gives an isomorphism from $\widehat{\mathbb{Z}(N)}$ to $\mathbb{Z}(N)$.

Example 2. The dual group of the circle[1] is precisely $\{e_n\}_{n\in\mathbb{Z}}$ (where $e_n(x) = e^{2\pi i n x}$). Moreover, $e_n \mapsto n$ gives an isomorphism between $\widehat{S^1}$ and the integers $\mathbb{Z}$.

Example 3. Characters on $\mathbb{R}$ are described by

$$e_\xi(x) = e^{2\pi i \xi x} \quad \text{where } \xi \in \mathbb{R}.$$

Thus $e_\xi \mapsto \xi$ is an isomorphism from $\widehat{\mathbb{R}}$ to $\mathbb{R}$.

Example 4. Since $\exp : \mathbb{R} \to \mathbb{R}^+$ is an isomorphism, we deduce from the previous example that the characters on $\mathbb{R}^+$ are given by

$$e_\xi(x) = x^{2\pi i \xi} = e^{2\pi i \xi \log x} \quad \text{where } \xi \in \mathbb{R},$$

and $\widehat{\mathbb{R}^+}$ is isomorphic to $\mathbb{R}$ (or $\mathbb{R}^+$).

The following lemma says that a nowhere vanishing multiplicative function is a character, a result that will be useful later.

Lemma 2.2 *Let G be a finite abelian group, and $e : G \to \mathbb{C} - \{0\}$ a multiplicative function, namely $e(a \cdot b) = e(a)e(b)$ for all $a, b \in G$. Then e is a character.*

[1] In addition to (2), the definition of a character on an infinite abelian group requires continuity. When G is the circle, $\mathbb{R}$, or $\mathbb{R}^+$, the meaning of "continuous" refers to the standard notion of limit.

Proof. The group G being finite, the absolute value of $e(a)$ is bounded above and below as a ranges over G. Since $|e(b^n)| = |e(b)|^n$, we conclude that $|e(b)| = 1$ for all $b \in G$.

The next step is to verify that the characters form an orthonormal basis of the vector space V of functions over the group G. This fact was obtained directly in the special case $G = \mathbb{Z}(N)$ from the explicit description of the characters $e_0, \dots, e_{N-1}$.

In the general case, we begin with the orthogonality relations; then we prove that there are "enough" characters by showing that there are as many as the order of the group.

2.3 The orthogonality relations

Let V denote the vector space of complex-valued functions defined on the finite abelian group G. Note that the dimension of V is $|G|$, the order of G. We define a Hermitian inner product on V by

$$(f, g) = \frac{1}{|G|} \sum_{a \in G} f(a)\overline{g(a)}, \qquad \text{whenever } f, g \in V. \tag{3}$$

Here the sum is taken over the group and is therefore finite.

Theorem 2.3 *The characters of G form an orthonormal family with respect to the inner product defined above.*

Since $|e(a)| = 1$ for any character, we find that

$$(e, e) = \frac{1}{|G|} \sum_{a \in G} e(a)\overline{e(a)} = \frac{1}{|G|} \sum_{a \in G} |e(a)|^2 = 1.$$

If $e \neq e'$ and both are characters, we must prove that $(e, e') = 0$; we isolate the key step in a lemma.

Lemma 2.4 *If e is a non-trivial character of the group G, then $\sum_{a \in G} e(a) = 0$.*

Proof. Choose $b \in G$ such that $e(b) \neq 1$. Then we have

$$e(b) \sum_{a \in G} e(a) = \sum_{a \in G} e(b)e(a) = \sum_{a \in G} e(ab) = \sum_{a \in G} e(a).$$

The last equality follows because as a ranges over the group, ab ranges over G as well. Therefore $\sum_{a \in G} e(a) = 0$.

We can now conclude the proof of the theorem. Suppose e' is a character distinct from e. Because $e(e')^{-1}$ is non-trivial, the lemma implies that

$$\sum_{a\in G} e(a)(e'(a))^{-1} = 0.$$

Since $(e'(a))^{-1} = \overline{e'(a)}$, the theorem is proved.

As a consequence of the theorem, we see that distinct characters are linearly independent. Since the dimension of V over $\mathbb{C}$ is $|G|$, we conclude that the order of $\hat{G}$ is finite and $\leq |G|$. The main result to which we now turn is that, in fact, $|\hat{G}| = |G|$.

2.4 Characters as a total family

The following completes the analogy between characters and the complex exponentials.

Theorem 2.5 *The characters of a finite abelian group G form a basis for the vector space of functions on G.*

There are several proofs of this theorem. One consists of using the structure theorem for finite abelian groups we have mentioned earlier, which states that any such group is the direct product of cyclic groups, that is, groups of the type $\mathbb{Z}(N)$. Since cyclic groups are self-dual, using this fact we would conclude that $|\hat{G}| = |G|$, and therefore the characters form a basis for G. (See Problem 3.)

Here we shall prove the theorem directly without these considerations.

Suppose V is a vector space of dimension d with inner product $(\cdot,\cdot)$. A linear transformation $T : V \to V$ is **unitary** if it preserves the inner product, $(Tv, Tw) = (v, w)$ for all $v, w \in V$. The spectral theorem from linear algebra asserts that any unitary transformation is diagonalizable. In other words, there exists a basis $\{v_1, \ldots, v_d\}$ (eigenvectors) of V such that $T(v_i) = \lambda_i v_i$, where $\lambda_i \in \mathbb{C}$ is the eigenvalue attached to v_i.

The proof of Theorem 2.5 is based on the following extension of the spectral theorem.

Lemma 2.6 *Suppose $\{T_1, \ldots, T_k\}$ is a commuting family of unitary transformations on the finite-dimensional inner product space V; that is,*

$$T_iT_j = T_jT_i \quad \text{for all } i, j.$$

Then $T_1, \ldots, T_k$ are simultaneously diagonalizable. In other words, there exists a basis for V which consists of eigenvectors for every T_i, $i = 1, \ldots, k$.

Proof. We use induction on k. The case $k = 1$ is simply the spectral theorem. Suppose that the lemma is true for any family of $k-1$ commuting unitary transformations. The spectral theorem applied to T_k says that V is the direct sum of its eigenspaces

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s},$$

where V_{λ_i} denotes the subspace of all eigenvectors with eigenvalue λ_i. We claim that each one of the $T_1, \ldots, T_{k-1}$ maps each eigenspace V_{λ_i} to itself. Indeed, if $v \in V_{\lambda_i}$ and $1 \leq j \leq k-1$, then

$$T_k T_j(v) = T_j T_k(v) = T_j(\lambda_i v) = \lambda_i T_j(v)$$

so $T_j(v) \in V_{\lambda_i}$, and the claim is proved.

Since the restrictions to V_{λ_i} of $T_1, \ldots, T_{k-1}$ form a family of commuting unitary linear transformations, the induction hypothesis guarantees that these are simultaneously diagonalizable on each subspace V_{λ_i}. This diagonalization provides us with the desired basis for each V_{λ_i}, and thus for V.

We can now prove Theorem 2.5. Recall that the vector space V of complex-valued functions defined on G has dimension $|G|$. For each $a \in G$ we define a linear transformation $T_a : V \to V$ by

$$(T_a f)(x) = f(a \cdot x) \quad \text{for } x \in G.$$

Since G is abelian it is clear that $T_a T_b = T_b T_a$ for all $a, b \in G$, and one checks easily that T_a is unitary for the Hermitian inner product (3) defined on V. By Lemma 2.6 the family $\{T_a\}_{a \in G}$ is simultaneously diagonalizable. This means there is a basis $\{v_b(x)\}_{b \in G}$ for V such that each $v_b(x)$ is an eigenfunction for T_a, for every a. Let v be one of these basis elements and 1 the unit element in G. We must have $v(1) \neq 0$ for otherwise

$$v(a) = v(a \cdot 1) = (T_a v)(1) = \lambda_a v(1) = 0,$$

where λ_a is the eigenvalue of v for T_a. Hence $v = 0$, and this is a contradiction. We claim that the function defined by $w(x) = \lambda_x = v(x)/v(1)$ is a character of G. Arguing as above we find that $w(x) \neq 0$ for every x, and

$$w(a \cdot b) = \frac{v(a \cdot b)}{v(1)} = \frac{\lambda_a v(b)}{v(1)} = \lambda_a \lambda_b \frac{v(1)}{v(1)} = \lambda_a \lambda_b = w(a)w(b).$$

We now invoke Lemma 2.2 to conclude the proof.

2.5 Fourier inversion and Plancherel formula

We now put together the results obtained in the previous sections to discuss the Fourier expansion of a function on a finite abelian group G. Given a function f on G and character e of G, we define the **Fourier coefficient** of f with respect to e, by

$$\hat{f}(e) = (f, e) = \frac{1}{|G|} \sum_{a \in G} f(a)\overline{e(a)},$$

and the **Fourier series** of f as

$$f \sim \sum_{e \in \hat{G}} \hat{f}(e) e.$$

Since the characters form a basis, we know that

$$f = \sum_{e \in \hat{G}} c_e e$$

for some set of constants c_e. By the orthogonality relations satisfied by the characters, we find that

$$(f, e) = c_e,$$

so f is indeed equal to its Fourier series, namely,

$$f = \sum_{e \in \hat{G}} \hat{f}(e) e.$$

We summarize our results.

Theorem 2.7 *Let G be a finite abelian group. The characters of G form an orthonormal basis for the vector space V of functions on G equipped with the inner product*

$$(f, g) = \frac{1}{|G|} \sum_{a \in G} f(a)\overline{g(a)}.$$

In particular, any function f on G is equal to its Fourier series

$$f = \sum_{e \in \hat{G}} \hat{f}(e) e.$$

Finally, we have the Parseval-Plancherel formula for finite abelian groups.

Theorem 2.8 *If f is a function on G, then* $\|f\|^2 = \sum_{e\in\hat{G}} |\hat{f}(e)|^2$.

Proof. Since the characters of G form an orthonormal basis for the vector space V, and $(f, e) = \hat{f}(e)$, we have that

$$\|f\|^2 = (f, f) = \sum_{e\in\hat{G}} (f, e)\overline{\hat{f}(e)} = \sum_{e\in\hat{G}} |\hat{f}(e)|^2.$$

The apparent difference of this statement with that of Theorem 1.2 is due to the different normalizations of the Fourier coefficients that are used.

3 Exercises

1. Let f be a function on the circle. For each $N \geq 1$ the discrete Fourier coefficients of f are defined by

$$a_N(n) = \frac{1}{N}\sum_{k=1}^{N} f(e^{2\pi ik/N})e^{-2\pi ikn/N}, \qquad \text{for } n \in \mathbb{Z}.$$

We also let

$$a(n) = \int_0^1 f(e^{2\pi ix})e^{-2\pi inx}\, dx$$

denote the ordinary Fourier coefficients of f.

(a) Show that $a_N(n) = a_N(n + N)$.

(b) Prove that if f is continuous, then $a_N(n) \to a(n)$ as $N \to \infty$.

2. If f is a C^1 function on the circle, prove that $|a_N(n)| \leq c/|n|$ whenever $0 < |n| \leq N/2$.
[Hint: Write

$$a_N(n)[1 - e^{2\pi i\ell n/N}] = \frac{1}{N}\sum_{k=1}^{N}[f(e^{2\pi ik/N}) - f(e^{2\pi i(k+\ell)/N})]e^{-2\pi ikn/N},$$

and choose ℓ so that $\ell n/N$ is nearly 1/2.]

3. By a similar method, show that if f is a C^2 function on the circle, then

$$|a_N(n)| \leq c/|n|^2, \qquad \text{whenever } 0 < |n| \leq N/2.$$

As a result, prove the inversion formula for $f \in C^2$,

$$f(e^{2\pi ix}) = \sum_{n=-\infty}^{\infty} a(n)e^{2\pi inx}$$

from its finite version.

[Hint: For the first part, use the second symmetric difference

$$f(e^{2\pi i(k+\ell)/N}) + f(e^{2\pi i(k-\ell)/N}) - 2f(e^{2\pi ik/N}).$$

For the second part, if N is odd (say), write the inversion formula as

$$f(e^{2\pi ik/N}) = \sum_{|n|<N/2} a_N(n)e^{2\pi ikn/N}.]$$

4. Let e be a character on $G = \mathbb{Z}(N)$, the additive group of integers modulo N. Show that there exists a unique $0 \leq \ell \leq N-1$ so that

$$e(k) = e_\ell(k) = e^{2\pi i\ell k/N} \qquad \text{for all } k \in \mathbb{Z}(N).$$

Conversely, every function of this type is a character on $\mathbb{Z}(N)$. Deduce that $e_\ell \mapsto \ell$ defines an isomorphism from $\hat{G}$ to G.

[Hint: Show that $e(1)$ is an Nth root of unity.]

5. Show that all characters on S^1 are given by

$$e_n(x) = e^{2\pi inx} \qquad \text{with } n \in \mathbb{Z},$$

and check that $e_n \mapsto n$ defines an isomorphism from $\widehat{S^1}$ to $\mathbb{Z}$.

[Hint: If F is continuous and $F(x+y) = F(x)F(y)$, then F is differentiable. To see this, note that if $F(0) \neq 0$, then for appropriate δ, $c = \int_0^\delta F(y)\,dy \neq 0$, and $cF(x) = \int_x^{\delta+x} F(y)\,dy$. Differentiate to conclude that $F(x) = e^{Ax}$ for some A.]

6. Prove that all characters on $\mathbb{R}$ take the form

$$e_\xi(x) = e^{2\pi i\xi x} \qquad \text{with } \xi \in \mathbb{R},$$

and that $e_\xi \mapsto \xi$ defines an isomorphism from $\widehat{\mathbb{R}}$ to $\mathbb{R}$. The argument in Exercise 5 applies here as well.

7. Let $\zeta = e^{2\pi i/N}$. Define the $N \times N$ matrix $M = (a_{jk})_{1\leq j,k\leq N}$ by $a_{jk} = N^{-1/2}\zeta^{jk}$.

(a) Show that M is unitary.

(b) Interpret the identity $(Mu, Mv) = (u, v)$ and the fact that $M^* = M^{-1}$ in terms of Fourier series on $\mathbb{Z}(N)$.

8. Suppose that $P(x) = \sum_{n=1}^{N} a_n e^{2\pi i n x}$.

(a) Show by using the Parseval identities for the circle and $\mathbb{Z}(N)$, that

$$\int_0^1 |P(x)|^2 \, dx = \frac{1}{N} \sum_{j=1}^{N} |P(j/N)|^2.$$

(b) Prove the reconstruction formula

$$P(x) = \sum_{j=1}^{N} P(j/N) K(x - (j/N))$$

where

$$K(x) = \frac{e^{2\pi i x}}{N} \frac{1 - e^{2\pi i N x}}{1 - e^{2\pi i x}} = \frac{1}{N}(e^{2\pi i x} + e^{2\pi i 2x} + \cdots + e^{2\pi i N x}).$$

Observe that P is completely determined by the values $P(j/N)$ for $1 \leq j \leq N$. Note also that $K(0) = 1$, and $K(j/N) = 0$ whenever j is not congruent to 0 modulo N.

9. To prove the following assertions, modify the argument given in the text.

(a) Show that one can compute the Fourier coefficients of a function on $\mathbb{Z}(N)$ when $N = 3^n$ with at most $6N \log_3 N$ operations.

(b) Generalize this to $N = \alpha^n$ where α is an integer > 1.

10. A group G is **cyclic** if there exists $g \in G$ that generates all of G, that is, if any element in G can be written as g^n for some $n \in \mathbb{Z}$. Prove that a finite abelian group is cyclic if and only if it is isomorphic to $\mathbb{Z}(N)$ for some N.

11. Write down the multiplicative tables for the groups $\mathbb{Z}^*(3)$, $\mathbb{Z}^*(4)$, $\mathbb{Z}^*(5)$, $\mathbb{Z}^*(6)$, $\mathbb{Z}^*(8)$, and $\mathbb{Z}^*(9)$. Which of these groups are cyclic?

12. Suppose that G is a finite abelian group and $e : G \to \mathbb{C}$ is a function that satisfies $e(x \cdot y) = e(x)e(y)$ for all $x, y \in G$. Prove that either e is identically 0, or e never vanishes. In the second case, show that for each x, $e(x) = e^{2\pi i r}$ for some $r \in \mathbb{Q}$ of the form $r = p/q$, where $q = |G|$.

13. In analogy with ordinary Fourier series, one may interpret finite Fourier expansions using convolutions as follows. Suppose G is a finite abelian group, 1_G its unit, and V the vector space of complex-valued functions on G.

(a) The convolution of two functions f and g in V is defined for each $a \in G$ by

$$(f * g)(a) = \frac{1}{|G|} \sum_{b \in G} f(b) g(a \cdot b^{-1}).$$

Show that for all $e \in \hat{G}$ one has $\widehat{(f * g)}(e) = \hat{f}(e)\hat{g}(e)$.

(b) Use Theorem 2.5 to show that if e is a character on G, then

$$\sum_{e \in \hat{G}} e(c) = 0 \qquad \text{whenever } c \in G \text{ and } c \neq 1_G.$$

(c) As a result of (b), show that the Fourier series $Sf(a) = \sum_{e \in \hat{G}} \hat{f}(e) e(a)$ of a function $f \in V$ takes the form

$$Sf = f * D,$$

where D is defined by

$$D(c) = \sum_{e \in \hat{G}} e(c) = \begin{cases} |G| & \text{if } c = 1_G, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Since $f * D = f$, we recover the fact that $Sf = f$. Loosely speaking, D corresponds to a "Dirac delta function"; it has unit mass

$$\frac{1}{|G|} \sum_{c \in G} D(c) = 1,$$

and (4) says that this mass is concentrated at the unit element in G. Thus D has the same interpretation as the "limit" of a family of good kernels. (See Section 4, Chapter 2.)

Note. The function D reappears in the next chapter as $\delta_1(n)$.

4 Problems

1. Prove that if n and m are two positive integers that are relatively prime, then

$$\mathbb{Z}(nm) \approx \mathbb{Z}(n) \times \mathbb{Z}(m).$$

[Hint: Consider the map $\mathbb{Z}(nm) \to \mathbb{Z}(n) \times \mathbb{Z}(m)$ given by $k \mapsto (k \bmod n, k \bmod m)$, and use the fact that there exist integers x and y such that $xn + ym = 1$.]

2.* Every finite abelian group G is isomorphic to a direct product of cyclic groups. Here are two more precise formulations of this theorem.

- If $p_1, \ldots, p_s$ are the distinct primes appearing in the factorization of the order of G, then

  $$G \approx G(p_1) \times \cdots \times G(p_s),$$

  where each $G(p)$ is of the form $G(p) = \mathbb{Z}(p^{r_1}) \times \cdots \times \mathbb{Z}(p^{r_\ell})$, with $0 \leq r_1 \leq \cdots \leq r_\ell$ (this sequence of integers depends on p of course). This decomposition is unique.

- There exist unique integers $d_1, \ldots, d_k$ such that

  $$d_1|d_2, \quad d_2|d_3, \quad \cdots, \quad d_{k-1}|d_k$$

  and

  $$G \approx \mathbb{Z}(d_1) \times \cdots \times \mathbb{Z}(d_k).$$

Deduce the second formulation from the first.

3. Let $\hat{G}$ denote the collection of distinct characters of the finite abelian group G.

(a) Note that if $G = \mathbb{Z}(N)$, then $\hat{G}$ is isomorphic to G.

(b) Prove that $\widehat{G_1 \times G_2} = \hat{G}_1 \times \hat{G}_2$.

(c) Prove using Problem 2 that if G is a finite abelian group, then $\hat{G}$ is isomorphic to G.

4.* When p is prime the group $\mathbb{Z}^*(p)$ is cyclic, and $\mathbb{Z}^*(p) \approx \mathbb{Z}(p-1)$.

8 Dirichlet's Theorem

Dirichlet, Gustav Lejeune (Düren 1805-Göttingen 1859), German mathematician. He was a number theorist at heart. But, while studying in Paris, being a very likeable person, he was befriended by Fourier and other like-minded mathematicians, and he learned analysis from them. Thus equipped, he was able to lay the foundation for the application of Fourier analysis to (analytic) theory of numbers.

S. Bochner, 1966

As a striking application of the theory of finite Fourier series, we now prove Dirichlet's theorem on primes in arithmetic progression. This theorem states that if q and ℓ are positive integers with no common factor, then the progression

$$\ell,\ \ell+q,\ \ell+2q,\ \ell+3q,\ldots,\ell+kq,\ldots$$

contains infinitely many prime numbers. This change of subject matter that we undertake illustrates the wide applicability of ideas from Fourier analysis to various areas outside its seemingly narrower confines. In this particular case, it is the theory of Fourier series on the finite abelian group $\mathbb{Z}^*(q)$ that plays a key role in the solution of the problem.

1 A little elementary number theory

We begin by introducing the requisite background. This involves elementary ideas of divisibility of integers, and in particular properties regarding prime numbers. Here the basic fact, called the fundamental theorem of arithmetic, is that every integer is the product of primes in an essentially unique way.

1.1 The fundamental theorem of arithmetic

The following theorem is a mathematical formulation of long division.

Theorem 1.1 (Euclid's algorithm) *For any integers a and b with $b > 0$, there exist unique integers q and r with $0 \leq r < b$ such that*

$$a = qb + r.$$

Here q denotes the quotient of a by b, and r is the remainder, which is smaller than b.

Proof. First we prove the existence of q and r. Let S denote the set of all non-negative integers of the form $a - qb$ with $q \in \mathbb{Z}$. This set is non-empty and in fact S contains arbitrarily large positive integers since $b \neq 0$. Let r denote the smallest element in S, so that

$$r = a - qb$$

for some integer q. By construction $0 \leq r$, and we claim that $r < b$. If not, we may write $r = b + s$ with $0 \leq s < r$, so $b + s = a - qb$, which then implies

$$s = a - (q + 1)b.$$

Hence $s \in S$ with $s < r$, and this contradicts the minimality of r. So $r < b$, hence q and r satisfy the conditions of the theorem.

To prove uniqueness, suppose we also had $a = q_1 b + r_1$ where $0 \leq r_1 < b$. By subtraction we find

$$(q - q_1)b = r_1 - r.$$

The left-hand side has absolute value 0 or $\geq b$, while the right-hand side has absolute value $< b$. Hence both sides of the equation must be 0, which gives $q = q_1$ and $r = r_1$.

An integer a **divides** b if there exists another integer c such that $ac = b$; we then write $a|b$ and say that a is a **divisor** of b. Note that in particular 1 divides every integer, and $a|a$ for all integers a. A **prime number** is a positive integer greater than 1 that has no positive divisors besides 1 and itself. The main theorem in this section says that any positive integer can be written uniquely as the product of prime numbers.

The **greatest common divisor** of two positive integers a and b is the largest integer that divides both a and b. We usually denote the greatest common divisor by $\gcd(a, b)$. Two positive integers are **relatively prime** if their greatest common divisor is 1. In other words, 1 is the only positive divisor common to both a and b.

Theorem 1.2 *If* $\gcd(a, b) = d$*, then there exist integers* x *and* y *such that*

$$ax + by = d.$$

Proof. Consider the set S of all positive integers of the form $ax + by$ where $x, y \in \mathbb{Z}$, and let s be the smallest element in S. We claim that $s = d$. By construction, there exist integers x and y such that

$$ax + by = s.$$

Clearly, any divisor of a and b divides s, so we must have $d \leq s$. The proof will be complete if we can show that $s|a$ and $s|b$. By Euclid's algorithm, we can write $a = qs + r$ with $0 \leq r < s$. Multiplying the above by q we find $qax + qby = qs$, and therefore

$$qax + qby = a - r.$$

Hence $r = a(1 - qx) + b(-qy)$. Since s was minimal in S and $0 \leq r < s$, we conclude that $r = 0$, therefore s divides a. A similar argument shows that s divides b, hence $s = d$ as desired.

In particular we record the following three consequences of the theorem.

Corollary 1.3 *Two positive integers* a *and* b *are relatively prime if and only if there exist integers* x *and* y *such that* $ax + by = 1$.

Proof. If a and b are relatively prime, two integers x and y with the desired property exist by Theorem 1.2. Conversely, if $ax + by = 1$ holds and d is positive and divides both a and b, then d divides 1, hence $d = 1$.

Corollary 1.4 *If* a *and* c *are relatively prime and* c *divides* ab*, then* c *divides* b*. In particular, if* p *is a prime that does not divide* a *and* p *divides* ab*, then* p *divides* b.

Proof. We can write $1 = ax + cy$, so multiplying by b we find $b = abx + cby$. Hence $c|b$.

Corollary 1.5 *If* p *is prime and* p *divides the product* $a_1 \cdots a_r$*, then* p *divides* a_i *for some* i.

Proof. By the previous corollary, if p does not divide a_1, then p divides $a_2 \cdots a_r$, so eventually $p|a_i$.

We can now prove the main result of this section.

Theorem 1.6 *Every positive integer greater than 1 can be factored uniquely into a product of primes.*

Proof. First, we show that such a factorization is possible. We do so by proving that the set S of positive integers > 1 which do not have a factorization into primes is empty. Arguing by contradiction, we assume that $S \neq \emptyset$. Let n be the smallest element of S. Since n cannot be a prime, there exist integers $a > 1$ and $b > 1$ such that $ab = n$. But then $a < n$ and $b < n$, so $a \notin S$ as well as $b \notin S$. Hence both a and b have prime factorizations and so does their product n. This implies $n \notin S$, therefore S is empty, as desired.

We now turn our attention to the uniqueness of the factorization. Suppose that n has two factorizations into primes

$$\begin{aligned} n &= p_1 p_2 \cdots p_r \\ &= q_1 q_2 \cdots q_s. \end{aligned}$$

So p_1 divides $q_1 q_2 \cdots q_s$, and we can apply Corollary 1.5 to conclude that $p_1 | q_i$ for some i. Since q_i is prime, we must have $p_1 = q_i$. Continuing with this argument we find that the two factorizations of n are equal up to a permutation of the factors.

We briefly digress to give an alternate definition of the group $\mathbb{Z}^*(q)$ which appeared in the previous chapter. According to our initial definition, $\mathbb{Z}^*(q)$ is the multiplicative group of units in $\mathbb{Z}(q)$: those $n \in \mathbb{Z}(q)$ for which there exists an integer m so that

$$nm \equiv 1 \mod q. \tag{1}$$

Equivalently, $\mathbb{Z}^*(q)$ is the group under multiplication of all integers in $\mathbb{Z}(q)$ that are relatively prime to q. Indeed, notice that if (1) is satisfied, then automatically n and q are relatively prime. Conversely, suppose we assume that n and q are relatively prime. Then, if we put $a = n$ and $b = q$ in Corollary 1.3, we find

$$nx + qy = 1.$$

Hence $nx \equiv 1 \mod q$, and we can take $m = x$ to establish the equivalence.

1.2 The infinitude of primes

The study of prime numbers has always been a central topic in arithmetic, and the first fundamental problem that arose was to determine whether

there are infinitely many primes or not. This problem was solved in Euclid's *Elements* with a simple and very elegant argument.

Theorem 1.7 *There are infinitely many primes.*

Proof. Suppose not, and denote by $p_1, \ldots, p_n$ the complete set of primes. Define

$$N = p_1 p_2 \cdots p_n + 1.$$

Since N is larger than any p_i, the integer N cannot be prime. Therefore, N is divisible by a prime that belongs to our list. But this is also an absurdity since every prime divides the product, yet no prime divides 1. This contradiction concludes the proof.

Euclid's argument actually can be modified to deduce finer results about the infinitude of primes. To see this, consider the following problem. Prime numbers (except for 2) can be divided into two classes depending on whether they are of the form $4k + 1$ or $4k + 3$, and the above theorem says that at least one of these classes has to be infinite. A natural question is to ask whether both classes are infinite, and if not, which one is? In the case of primes of the form $4k + 3$, the fact that the class is infinite has a proof that is similar to Euclid's, but with a twist. If there are only finitely many such primes, enumerate them in increasing order omitting 3,

$$p_1 = 7, \quad p_2 = 11, \ \ldots, \ p_n,$$

and let

$$N = 4p_1 p_2 \cdots p_n + 3.$$

Clearly, N is of the form $4k + 3$ and cannot be prime since $N > p_n$. Since the product of two numbers of the form $4m + 1$ is again of the form $4m + 1$, one of the prime divisors of N, say p, must be of the form $4k + 3$. We must have $p \neq 3$, since 3 does not divide the product in the definition of N. Also, p cannot be one of the other primes of the form $4k + 3$, that is, $p \neq p_i$ for $i = 1, \ldots n$, because then p divides the product $p_1 \cdots p_n$ but does not divide 3.

It remains to determine if the class of primes of the form $4k + 1$ is infinite. A simple-minded modification of the above argument does not work since the product of two numbers of the form $4m + 3$ is never of the form $4m + 3$. More generally, in an attempt to prove the law of quadratic reciprocity, Legendre formulated the following statement:

If q and ℓ are relatively prime, then the sequence

$$\ell + kq, \qquad k \in \mathbb{Z}$$

contains infinitely many primes (hence at least one prime!).

Of course, the condition that q and ℓ be relatively prime is necessary, for otherwise $\ell + kq$ is never prime. In other words, this hypothesis says that any arithmetic progression that could contain primes necessarily contains infinitely many of them.

Legendre's assertion was proved by Dirichlet. The key idea in his proof is Euler's analytical approach to prime numbers involving his product formula, which gives a strengthened version of Theorem 1.7. This insight of Euler led to a deep connection between the theory of primes and analysis.

The zeta function and its Euler product

We begin with a rapid review of infinite products. If $\{A_n\}_{n=1}^{\infty}$ is a sequence of real numbers, we define

$$\prod_{n=1}^{\infty} A_n = \lim_{N\to\infty} \prod_{n=1}^{N} A_n$$

if the limit exists, in which case we say that the product converges. The natural approach is to take logarithms and transform products into sums. We gather in a lemma the properties we shall need of the function $\log x$, defined for positive real numbers.

Lemma 1.8 *The exponential and logarithm functions satisfy the following properties:*

(i) $e^{\log x} = x$.

(ii) $\log(1+x) = x + E(x)$ *where* $|E(x)| \le x^2$ *if* $|x| < 1/2$.

(iii) *If* $\log(1+x) = y$ *and* $|x| < 1/2$, *then* $|y| \le 2|x|$.

In terms of the O notation, property (ii) will be recorded as $\log(1+x) = x + O(x^2)$.

Proof. Property (i) is standard. To prove property (ii) we use the power series expansion of $\log(1+x)$ for $|x| < 1$, that is,

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n. \tag{2}$$

Then we have

$$E(x) = \log(1+x) - x = -\frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots ,$$

and the triangle inequality implies

$$|E(x)| \le \frac{x^2}{2}\left(1 + |x| + |x|^2 + \cdots\right).$$

Therefore, if $|x| \le 1/2$ we can sum the geometric series on the right-hand side to find that

$$\begin{aligned} |E(x)| &\le \frac{x^2}{2}\left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) \\ &\le \frac{x^2}{2}\left(\frac{1}{1-1/2}\right) \\ &\le x^2. \end{aligned}$$

The proof of property (iii) is now immediate; if $x \neq 0$ and $|x| \le 1/2$, then

$$\begin{aligned} \left|\frac{\log(1+x)}{x}\right| &\le 1 + \left|\frac{E(x)}{x}\right| \\ &\le 1 + |x| \\ &\le 2, \end{aligned}$$

and if $x = 0$, (iii) is clearly also true.

We can now prove the main result on infinite products of real numbers.

Proposition 1.9 *If $A_n = 1 + a_n$ and $\sum |a_n|$ converges, then the product $\prod_n A_n$ converges, and this product vanishes if and only if one of its factors A_n vanishes. Also, if $a_n \neq 1$ for all n, then $\prod_n 1/(1-a_n)$ converges.*

Proof. If $\sum |a_n|$ converges, then for all large n we must have $|a_n| < 1/2$. Disregarding finitely many terms if necessary, we may assume that this inequality holds for all n. Then we may write the partial products as follows:

$$\prod_{n=1}^{N} A_n = \prod_{n=1}^{N} e^{\log(1+a_n)} = e^{B_N},$$

where $B_N = \sum_{n=1}^{N} b_n$ with $b_n = \log(1 + a_n)$. By the lemma, we know that $|b_n| \le 2|a_n|$, so that B_N converges to a real number, say B. Since

the exponential function is continuous, we conclude that e^{B_N} converges to e^B as N goes to infinity, proving the first assertion of the proposition. Observe also that if $1 + a_n \neq 0$ for all n, the product converges to a non-zero limit since it is expressed as e^B.

Finally observe that the partial products of $\prod_n 1/(1 - a_n)$ are $1/\prod_{n=1}^N (1 - a_n)$, so the same argument as above proves that the product in the denominator converges to a non-zero limit.

With these preliminaries behind us, we can now return to the heart of the matter. For s a real number (strictly) greater than 1, we define the **zeta function** by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

To see that the series defining ζ converges, we use the principle that whenever f is a decreasing function one can compare $\sum f(n)$ with $\int f(x)\, dx$, as is suggested by Figure 1. Note also that a similar technique was used in Chapter 3, that time bounding a sum from below by an integral.

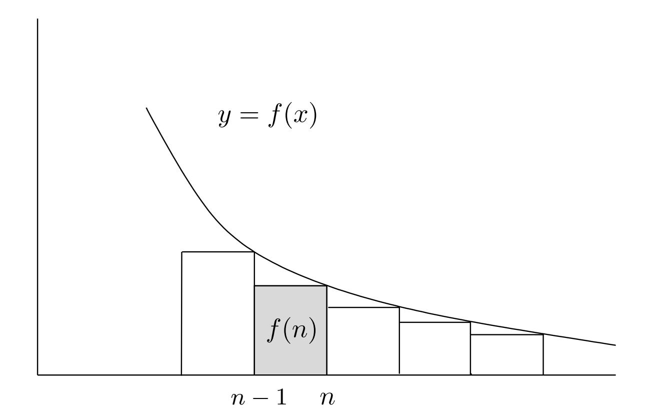


Figure 1. Comparing sums with integrals

Here we take $f(x) = 1/x^s$ to see that

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \leq 1 + \sum_{n=2}^{\infty} \int_{n-1}^{n} \frac{dx}{x^s} = 1 + \int_1^{\infty} \frac{dx}{x^s},$$

and therefore,

$$\zeta(s) \leq 1 + \frac{1}{s-1}. \tag{3}$$

Clearly, the series defining ζ converges uniformly on each half-line $s > s_0 > 1$, hence ζ is continuous when $s > 1$. The zeta function was already mentioned earlier in the discussion of the Poisson summation formula and the theta function.

The key result is Euler's product formula.

Theorem 1.10 *For every $s > 1$, we have*

$$\zeta(s) = \prod_p \frac{1}{1 - 1/p^s},$$

where the product is taken over all primes.

It is important to remark that this identity is an analytic expression of the fundamental theorem of arithmetic. In fact, each factor of the product $1/(1-p^{-s})$ can be written as a convergent geometric series

$$1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{Ms}} + \cdots .$$

So we consider

$$\prod_{p_j} \left(1 + \frac{1}{p_j^s} + \frac{1}{p_j^{2s}} + \cdots + \frac{1}{p_j^{Ms}} + \cdots \right),$$

where the product is taken over all primes, which we order in increasing order $p_1 < p_2 < \cdots$. Proceeding formally (these manipulations will be justified below), we calculate the product as a sum of terms, each term originating by picking out a term $1/p_j^{ks}$ (in the sum corresponding to p_j) with a k, which of course will depend on j, and with $k = 0$ for j sufficiently large. The product obtained this way is

$$\frac{1}{(p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m})^s} = \frac{1}{n^s},$$

where the integer n is written as a product of primes $n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$. By the fundamental theorem of arithmetic, each integer ≥ 1 occurs in this way uniquely, hence the product equals

$$\sum_{n=1}^{\infty} \frac{1}{n^s}.$$

We now justify this heuristic argument.

Proof. Suppose M and N are positive integers with $M > N$. Observe now that any positive integer $n \leq N$ can be written uniquely as a product of primes, and that each prime must be less than or equal to N and repeated less than M times. Therefore

$$\begin{aligned}\sum_{n=1}^{N} \frac{1}{n^s} &\leq \prod_{p \leq N} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{Ms}}\right) \\ &\leq \prod_{p \leq N} \left(\frac{1}{1 - p^{-s}}\right) \\ &\leq \prod_{p} \left(\frac{1}{1 - p^{-s}}\right).\end{aligned}$$

Letting N tend to infinity now yields

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \leq \prod_{p} \left(\frac{1}{1 - p^{-s}}\right).$$

For the reverse inequality, we argue as follows. Again, by the fundamental theorem of arithmetic, we find that

$$\prod_{p \leq N} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{Ms}}\right) \leq \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Letting M tend to infinity gives

$$\prod_{p \leq N} \left(\frac{1}{1 - p^{-s}}\right) \leq \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Hence

$$\prod_{p} \left(\frac{1}{1 - p^{-s}}\right) \leq \sum_{n=1}^{\infty} \frac{1}{n^s},$$

and the proof of the product formula is complete.

We now come to Euler's version of Theorem 1.7, which inspired Dirichlet's approach to the general problem of primes in arithmetic progression. The point is the following proposition.

Proposition 1.11 *The series*

$$\sum_p 1/p$$

diverges, when the sum is taken over all primes p.

Of course, if there were only finitely many primes the series would converge automatically.

Proof. We take logarithms of both sides of the Euler formula. Since $\log x$ is continuous, we may write the logarithm of the infinite product as the sum of the logarithms. Therefore, we obtain for $s > 1$

$$-\sum_p \log(1 - 1/p^s) = \log \zeta(s).$$

Since $\log(1 + x) = x + O(|x|^2)$ whenever $|x| \leq 1/2$, we get

$$-\sum_p \left[-1/p^s + O(1/p^{2s})\right] = \log \zeta(s),$$

which gives

$$\sum_p 1/p^s + O(1) = \log \zeta(s).$$

The term $O(1)$ appears because $\sum_p 1/p^{2s} \leq \sum_{n=1}^{\infty} 1/n^2$. Now we let s tend to 1 from above, namely $s \to 1^+$, and note that $\zeta(s) \to \infty$ since $\sum_{n=1}^{\infty} 1/n^s \geq \sum_{n=1}^{M} 1/n^s$, and therefore

$$\liminf_{s \to 1^+} \sum_{n=1}^{\infty} 1/n^s \geq \sum_{n=1}^{M} 1/n \qquad \text{for every } M.$$

We conclude that $\sum_p 1/p^s \to \infty$ as $s \to 1^+$, and since $1/p > 1/p^s$ for all $s > 1$, we finally have that

$$\sum_p 1/p = \infty.$$

In the rest of this chapter we see how Dirichlet adapted Euler's insight.

2 Dirichlet's theorem

We remind the reader of our goal:

Theorem 2.1 *If q and ℓ are relatively prime positive integers, then there are infinitely many primes of the form $\ell + kq$ with $k \in \mathbb{Z}$.*

Following Euler's argument, Dirichlet proved this theorem by showing that the series

$$\sum_{p \equiv \ell \mod q} \frac{1}{p}$$

diverges, where the sum is over all primes congruent to ℓ modulo q. Once q is fixed and no confusion is possible, we write $p \equiv \ell$ to denote a prime congruent to ℓ modulo q. The proof consists of several steps, one of which requires Fourier analysis on the group $\mathbb{Z}^*(q)$. Before proceeding with the theorem in its complete generality, we outline the solution to the particular problem raised earlier: are there infinitely many primes of the form $4k+1$? This example, which consists of the special case $q = 4$ and $\ell = 1$, illustrates all the important steps in the proof of Dirichlet's theorem.

We begin with the character on $\mathbb{Z}^*(4)$ defined by $\chi(1) = 1$ and $\chi(3) = -1$. We extend this character to all of $\mathbb{Z}$ as follows:

$$\chi(n) = \begin{cases} 0 & \text{if } n \text{ is even,} \\ 1 & \text{if } n = 4k+1, \\ -1 & \text{if } n = 4k+3. \end{cases}$$

Note that this function is multiplicative, that is, $\chi(nm) = \chi(n)\chi(m)$ on *all* of $\mathbb{Z}$. Let $L(s, \chi) = \sum_{n=1}^{\infty} \chi(n)/n^s$, so that

$$L(s, \chi) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots .$$

Then $L(1, \chi)$ is the convergent series given by

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots .$$

Since the terms in the series are alternating and their absolute values decrease to zero we have $L(1, \chi) \neq 0$. Because χ is multiplicative, the Euler product generalizes (as we will prove later) to give

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)/p^s}.$$

Taking the logarithm of both sides, we find that

$$\log L(s,\chi) = \sum_p \frac{\chi(p)}{p^s} + O(1).$$

Letting $s \to 1^+$, the observation that $L(1,\chi) \neq 0$ shows that $\sum_p \chi(p)/p^s$ remains bounded. Hence

$$\sum_{p\equiv 1} \frac{1}{p^s} - \sum_{p\equiv 3} \frac{1}{p^s}$$

is bounded as $s \to 1^+$. However, we know from Proposition 1.11 that

$$\sum_p \frac{1}{p^s}$$

is unbounded as $s \to 1^+$, so putting these two facts together, we find that

$$2\sum_{p\equiv 1} \frac{1}{p^s}$$

is unbounded as $s \to 1^+$. Hence $\sum_{p\equiv 1} 1/p$ diverges, and as a consequence there are infinitely many primes of the form $4k+1$.

We digress briefly to show that in fact $L(1,\chi) = \pi/4$. To see this, we integrate the identity

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots ,$$

and get

$$\int_0^y \frac{dx}{1+x^2} = y - \frac{y^3}{3} + \frac{y^5}{5} - \cdots , \qquad 0 < y < 1.$$

We then let y tend to 1. The integral can be calculated as

$$\int_0^1 \frac{dx}{1+x^2} = \arctan u\big|_0^1 = \frac{\pi}{4},$$

so this proves that the series $1 - 1/3 + 1/5 - \cdots$ is Abel summable to $\pi/4$. Since we know the series converges, its limit is the same as its Abel limit, hence $1 - 1/3 + 1/5 - \cdots = \pi/4$.

The rest of this chapter gives the full proof of Dirichlet's theorem. We begin with the Fourier analysis (which is actually the last step in the example given above), and reduce the theorem to the non-vanishing of L-functions.

2.1 Fourier analysis, Dirichlet characters, and reduction of the theorem

In what follows we take the abelian group G to be $\mathbb{Z}^*(q)$. Our formulas below involve the order of G, which is the number of integers $0 \leq n < q$ that are relatively prime to q; this number defines the **Euler phi-function** $\varphi(q)$, and $|G| = \varphi(q)$.

Consider the function δ_ℓ on G, which we think of as the characteristic function of ℓ; if $n \in \mathbb{Z}^*(q)$, then

$$\delta_\ell(n) = \begin{cases} 1 & \text{if } n \equiv \ell \mod q, \\ 0 & \text{otherwise.} \end{cases}$$

We can expand this function in a Fourier series as follows:

$$\delta_\ell(n) = \sum_{e \in \hat{G}} \widehat{\delta_\ell}(e) e(n),$$

where

$$\widehat{\delta_\ell}(e) = \frac{1}{|G|} \sum_{m \in G} \delta_\ell(m) \overline{e(m)} = \frac{1}{|G|} \overline{e(\ell)}.$$

Hence

$$\delta_\ell(n) = \frac{1}{|G|} \sum_{e \in \hat{G}} \overline{e(\ell)} e(n).$$

We can extend the function δ_ℓ to all of $\mathbb{Z}$ by setting $\delta_\ell(m) = 0$ whenever m and q are not relatively prime. Similarly, the extensions of the characters $e \in \hat{G}$ to all of $\mathbb{Z}$ which are given by the recipe

$$\chi(m) = \begin{cases} e(m) & \text{if } m \text{ and } q \text{ are relatively prime} \\ 0 & \text{otherwise,} \end{cases}$$

are called the **Dirichlet characters** modulo q. We shall denote the extension to $\mathbb{Z}$ of the trivial character of G by χ_0, so that $\chi_0(m) = 1$ if m and q are relatively prime, and 0 otherwise. Note that the Dirichlet characters modulo q are multiplicative on all of $\mathbb{Z}$, in the sense that

$$\chi(nm) = \chi(n)\chi(m) \quad \text{for all } n, m \in \mathbb{Z}.$$

Since the integer q is fixed, we may without fear of confusion, speak of "Dirichlet characters" omitting reference to q.[1]

With $|G| = \varphi(q)$, we may restate the above results as follows:

---

[1] We use the notation χ instead of e to distinguish the Dirichlet characters (defined on $\mathbb{Z}$) from the characters e (defined on $\mathbb{Z}^*(q)$).

Lemma 2.2 *The Dirichlet characters are multiplicative. Moreover,*

$$\delta_\ell(m) = \frac{1}{\varphi(q)} \sum_\chi \overline{\chi(\ell)}\chi(m),$$

where the sum is over all Dirichlet characters.

With the above lemma we have taken our first step towards a proof of the theorem, since this lemma shows that

$$\begin{aligned}\sum_{p\equiv\ell} \frac{1}{p^s} &= \sum_p \frac{\delta_\ell(p)}{p^s} \\ &= \frac{1}{\varphi(q)} \sum_\chi \overline{\chi(\ell)} \sum_p \frac{\chi(p)}{p^s}.\end{aligned}$$

Thus it suffices to understand the behavior of $\sum_p \chi(p)p^{-s}$ as $s \to 1^+$. In fact, we divide the above sum in two parts depending on whether or not χ is trivial. So we have

$$\begin{aligned}\sum_{p\equiv\ell} \frac{1}{p^s} &= \frac{1}{\varphi(q)} \sum_p \frac{\chi_0(p)}{p^s} + \frac{1}{\varphi(q)} \sum_{\chi\neq\chi_0} \overline{\chi(\ell)} \sum_p \frac{\chi(p)}{p^s} \\ &= \frac{1}{\varphi(q)} \sum_{p \text{ not dividing } q} \frac{1}{p^s} + \frac{1}{\varphi(q)} \sum_{\chi\neq\chi_0} \overline{\chi(\ell)} \sum_p \frac{\chi(p)}{p^s}.\end{aligned} \tag{4}$$

Since there are only finitely many primes dividing q, Euler's theorem (Proposition 1.11) implies that the first sum on the right-hand side diverges when s tends to 1. These observations show that Dirichlet's theorem is a consequence of the following assertion.

Theorem 2.3 *If χ is a nontrivial Dirichlet character, then the sum*

$$\sum_p \frac{\chi(p)}{p^s}$$

remains bounded as $s \to 1^+$.

The proof of Theorem 2.3 requires the introduction of the L-functions, to which we now turn.

2.2 Dirichlet L-functions

We proved earlier that the zeta function $\zeta(s) = \sum_n 1/n^s$ could be expressed as a product, namely

$$\sum_{n=1}^\infty \frac{1}{n^s} = \prod_p \frac{1}{(1-p^{-s})}.$$

Dirichlet observed an analogue of this formula for the so-called **L-functions** defined for $s > 1$ by

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

where χ is a Dirichlet character.

Theorem 2.4 *If $s > 1$, then*

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{(1 - \chi(p)p^{-s})},$$

where the product is over all primes.

Assuming this theorem for now, we can follow Euler's argument formally: taking the logarithm of the product and using the fact that $\log(1 + x) = x + O(x^2)$ whenever x is small, we would get

$$\begin{aligned}\log L(s, \chi) &= -\sum_p \log(1 - \chi(p)/p^s) \\ &= -\sum_p \left[-\frac{\chi(p)}{p^s} + O\left(\frac{1}{p^{2s}}\right)\right] \\ &= \sum_p \frac{\chi(p)}{p^s} + O(1).\end{aligned}$$

If $L(1, \chi)$ is finite and non-zero, then $\log L(s, \chi)$ is bounded as $s \to 1^+$, and we can conclude that the sum

$$\sum_p \frac{\chi(p)}{p^s}$$

is bounded as $s \to 1^+$. We now make several observations about the above formal argument.

First, we must prove the product formula in Theorem 2.4. Since the Dirichlet characters χ can be complex-valued we will extend the logarithm to complex numbers w of the form $w = 1/(1 - z)$ with $|z| < 1$. (This will be done in terms of a power series.) Then we show that with this definition of the logarithm, the proof of Euler's product formula given earlier carries over to L-functions.

Second, we must make sense of taking the logarithm of both sides of the product formula. If the Dirichlet characters are real, this argument works

and is precisely the one given in the example corresponding to primes of the form $4k+1$. In general, the difficulty lies in the fact that $\chi(p)$ is a complex number, and the complex logarithm is not single valued; in particular, the logarithm of a product is not the sum of the logarithms.

Third, it remains to prove that whenever $\chi \neq \chi_0$, then $\log L(s,\chi)$ is bounded as $s \to 1^+$. If (as we shall see) $L(s,\chi)$ is continuous at $s=1$, then it suffices to show that

$$L(1,\chi) \neq 0.$$

This is the non-vanishing we mentioned earlier, which corresponds to the alternating series being non-zero in the previous example. The fact that $L(1,\chi) \neq 0$ is the most difficult part of the argument.

So we will focus on three points:

1. Complex logarithms and infinite products.
2. Study of $L(s,\chi)$.
3. Proof that $L(1,\chi) \neq 0$ if χ is non-trivial.

However, before we enter further into the details, we pause briefly to discuss some historical facts surrounding Dirichlet's theorem.

Historical digression

In the following list, we have gathered the names of those mathematicians whose work dealt most closely with the series of achievements related to Dirichlet's theorem. To give a better perspective, we attach the years in which they reached the age of 35:

Euler 1742
Legendre 1787
Gauss 1812
Dirichlet 1840
Riemann 1861

As we mentioned earlier, Euler's discovery of the product formula for the zeta function is the starting point in Dirichlet's argument. Legendre in effect conjectured the theorem because he needed it in his proof of the law of quadratic reciprocity. However, this goal was first accomplished by Gauss who, while not knowing how to establish the theorem about primes in arithmetic progression, nevertheless found a number of different proofs of quadratic reciprocity. Later, Riemann extended the study of the zeta function to the complex plane and indicated how properties

related to the non-vanishing of that function were central in the further understanding of the distribution of prime numbers.

Dirichlet proved his theorem in 1837. It should be noted that Fourier, who had befriended Dirichlet when the latter was a young mathematician visiting Paris, had died several years before. Besides the great activity in mathematics, that period was also a very fertile time in the arts, and in particular music. The era of Beethoven had ended only ten years earlier, and Schumann was now reaching the heights of his creativity. But the musician whose career was closest to Dirichlet was Felix Mendelssohn (four years his junior). It so happens that the latter began composing his famous violin concerto the year after Dirichlet succeeded in proving his theorem.

3 Proof of the theorem

We return to the proof of Dirichlet's theorem and to the three difficulties mentioned above.

3.1 Logarithms

The device to deal with the first point is to define two logarithms, one for complex numbers of the form $1/(1-z)$ with $|z|<1$ which we denote by $\log_1$, and one for the function $L(s,\chi)$ which we will denote by $\log_2$.

For the first logarithm, we define

$$\log_1\left(\frac{1}{1-z}\right)=\sum_{k=1}^{\infty}\frac{z^k}{k}\qquad\text{for } |z|<1.$$

Note that $\log_1 w$ is then defined if $\text{Re}(w)>1/2$, and because of equation (2), $\log_1 w$ gives an extension of the usual $\log x$ when x is a real number $>1/2$.

Proposition 3.1 *The logarithm function* $\log_1$ *satisfies the following properties:*

(i) *If* $|z|<1$, *then*

$$e^{\log_1\left(\frac{1}{1-z}\right)}=\frac{1}{1-z}.$$

(ii) *If* $|z|<1$, *then*

$$\log_1\left(\frac{1}{1-z}\right)=z+E_1(z),$$

where the error E_1 *satisfies* $|E_1(z)|\leq|z|^2$ *if* $|z|<1/2$.

(iii) *If* $|z| < 1/2$, *then*

$$\left|\log_1\left(\frac{1}{1-z}\right)\right| \le 2|z|.$$

Proof. To establish the first property, let $z = re^{i\theta}$ with $0 \le r < 1$, and observe that it suffices to show that

$$(1 - re^{i\theta})\, e^{\sum_{k=1}^{\infty}(re^{i\theta})^k/k} = 1. \tag{5}$$

To do so, we differentiate the left-hand side with respect to r, and this gives

$$\left[-e^{i\theta} + (1 - re^{i\theta})\left(\sum_{k=1}^{\infty}(re^{i\theta})^k/k\right)'\right] e^{\sum_{k=1}^{\infty}(re^{i\theta})^k/k}.$$

The term in brackets equals

$$-e^{i\theta} + (1 - re^{i\theta})e^{i\theta}\left(\sum_{k=1}^{\infty}(re^{i\theta})^{k-1}\right) = -e^{i\theta} + (1 - re^{i\theta})e^{i\theta}\frac{1}{1 - re^{i\theta}} = 0.$$

Having found that the left-hand side of the equation (5) is constant, we set $r = 0$ and get the desired result.

The proofs of the second and third properties are the same as their real counterparts given in Lemma 1.8.

Using these results we can state a sufficient condition guaranteeing the convergence of infinite products of complex numbers. Its proof is the same as in the real case, except that we now use the logarithm $\log_1$.

Proposition 3.2 *If* $\sum |a_n|$ *converges, and* $a_n \ne 1$ *for all* n, *then*

$$\prod_{n=1}^{\infty}\left(\frac{1}{1-a_n}\right)$$

converges. Moreover, this product is non-zero.

Proof. For n large enough, $|a_n| < 1/2$, so we may assume without loss of generality that this inequality holds for all $n \ge 1$. Then

$$\prod_{n=1}^{N}\left(\frac{1}{1-a_n}\right) = \prod_{n=1}^{N} e^{\log_1\left(\frac{1}{1-a_n}\right)} = e^{\sum_{n=1}^{N}\log_1\left(\frac{1}{1-a_n}\right)}.$$

But we know from the previous proposition that

$$\left|\log_1\left(\frac{1}{1-z}\right)\right| \le 2|z|,$$

so the fact that the series $\sum |a_n|$ converges, immediately implies that the limit

$$\lim_{N\to\infty}\sum_{n=1}^{N}\log_1\left(\frac{1}{1-a_n}\right) = A$$

exists. Since the exponential function is continuous, we conclude that the product converges to e^A, which is clearly non-zero.

We may now prove the promised Dirichlet product formula

$$\sum_n \frac{\chi(n)}{n^s} = \prod_p \frac{1}{(1-\chi(p)p^{-s})}.$$

For simplicity of notation, let L denote the left-hand side of the above equation. Define

$$S_N = \sum_{n\le N}\chi(n)n^{-s} \quad \text{and} \quad \Pi_N = \prod_{p\le N}\left(\frac{1}{1-\chi(p)p^{-s}}\right).$$

The infinite product $\Pi = \lim_{N\to\infty}\Pi_N = \prod_p\left(\frac{1}{1-\chi(p)p^{-s}}\right)$ converges by the previous proposition. Indeed, if we set $a_n = \chi(p_n)p_n^{-s}$, where p_n is the nth prime, we note that if $s > 1$, then $\sum |a_n| < \infty$.

Also, define

$$\Pi_{N,M} = \prod_{p\le N}\left(1 + \frac{\chi(p)}{p^s} + \cdots + \frac{\chi(p^M)}{p^{Ms}}\right).$$

Now fix $\epsilon > 0$ and choose N so large that

$$|S_N - L| < \epsilon \quad \text{and} \quad |\Pi_N - \Pi| < \epsilon.$$

We can next select M large enough so that

$$|S_N - \Pi_{N,M}| < \epsilon \quad \text{and} \quad |\Pi_{N,M} - \Pi_N| < \epsilon.$$

To see the first inequality, one uses the fundamental theorem of arithmetic and the fact that the Dirichlet characters are multiplicative. The

second inequality follows merely because each series $\sum_{n=1}^{\infty} \frac{\chi(p^n)}{p^{ns}}$ converges.

Therefore

$$|L - \Pi| \leq |L - S_N| + |S_N - \Pi_{N,M}| + |\Pi_{N,M} - \Pi_N| + |\Pi_N - \Pi| < 4\epsilon,$$

as was to be shown.

3.2 L-functions

The next step is a better understanding of the L-functions. Their behavior as functions of s (especially near $s = 1$) depends on whether or not χ is trivial. In the first case, $L(s, \chi_0)$ is up to some simple factors just the zeta function.

Proposition 3.3 *Suppose χ_0 is the trivial Dirichlet character,*

$$\chi_0(n) = \begin{cases} 1 & \text{if } n \text{ and } q \text{ are relatively prime,} \\ 0 & \text{otherwise,} \end{cases}$$

and $q = p_1^{a_1} \cdots p_N^{a_N}$ is the prime factorization of q. Then

$$L(s, \chi_0) = (1 - p_1^{-s})(1 - p_2^{-s}) \cdots (1 - p_N^{-s})\zeta(s).$$

Therefore $L(s, \chi_0) \to \infty$ as $s \to 1^+$.

Proof. The identity follows at once on comparing the Dirichlet and Euler product formulas. The final statement holds because $\zeta(s) \to \infty$ as $s \to 1^+$.

The behavior of the remaining L-functions, those for which $\chi \neq \chi_0$, is more subtle. A remarkable property is that these functions are now defined and continuous for $s > 0$. In fact, more is true.

Proposition 3.4 *If χ is a non-trivial Dirichlet character, then the series*

$$\sum_{n=1}^{\infty} \chi(n)/n^s$$

converges for $s > 0$, and we denote its sum by $L(s, \chi)$. Moreover:

(i) *The function $L(s, \chi)$ is continuously differentiable for $0 < s < \infty$.*

(ii) *There exists constants $c, c' > 0$ so that*

$$L(s, \chi) = 1 + O(e^{-cs}) \quad \text{as } s \to \infty, \text{ and}$$

$$L'(s, \chi) = O(e^{-c's}) \qquad \text{as } s \to \infty.$$

We first isolate the key cancellation property that non-trivial Dirichlet characters possess, which accounts for the behavior of the L-function described in the proposition.

Lemma 3.5 *If χ is a non-trivial Dirichlet character, then*

$$\left|\sum_{n=1}^{k}\chi(n)\right| \leq q, \qquad \textit{for any } k.$$

Proof. First, we recall that

$$\sum_{n=1}^{q}\chi(n) = 0.$$

In fact, if S denotes the sum and $a \in \mathbb{Z}^*(q)$, then the multiplicative property of the Dirichlet character χ gives

$$\chi(a)S = \sum \chi(a)\chi(n) = \sum \chi(an) = \sum \chi(n) = S.$$

Since χ is non-trivial, $\chi(a) \neq 1$ for some a, hence $S = 0$. We now write $k = aq + b$ with $0 \leq b < q$, and note that

$$\sum_{n=1}^{k}\chi(n) = \sum_{n=1}^{aq}\chi(n) + \sum_{aq<n\leq aq+b}\chi(n) = \sum_{aq<n\leq aq+b}\chi(n),$$

and there are no more than q terms in the last sum. The proof is complete once we recall that $|\chi(n)| \leq 1$.

We can now prove the proposition. Let $s_k = \sum_{n=1}^{k}\chi(n)$, and $s_0 = 0$. We know that $L(s,\chi)$ is defined for $s > 1$ by the series

$$\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}$$

which converges absolutely and uniformly for $s > \delta > 1$. Moreover, the differentiated series also converges absolutely and uniformly for $s > \delta > 1$, which shows that $L(s,\chi)$ is continuously differentiable for $s > 1$. We

sum by parts[2] to extend this result to $s > 0$. Indeed, we have

$$\begin{aligned}\sum_{k=1}^{N} \frac{\chi(k)}{k^s} &= \sum_{k=1}^{N} \frac{s_k - s_{k-1}}{k^s} \\ &= \sum_{k=1}^{N-1} s_k \left[\frac{1}{k^s} - \frac{1}{(k+1)^s}\right] + \frac{s_N}{N^s} \\ &= \sum_{k=1}^{N-1} f_k(s) + \frac{s_N}{N^s},\end{aligned}$$

where $f_k(s) = s_k\,[k^{-s} - (k+1)^{-s}]$. If $g(x) = x^{-s}$, then $g'(x) = -sx^{-s-1}$, so applying the mean-value theorem between $x = k$ and $x = k+1$, and the fact that $|s_k| \leq q$, we find that

$$|f_k(s)| \leq qsk^{-s-1}.$$

Therefore, the series $\sum f_k(s)$ converges absolutely and uniformly for $s > \delta > 0$, and this proves that $L(s, \chi)$ is continuous for $s > 0$. To prove that it is also continuously differentiable, we differentiate the series term by term, obtaining

$$\sum (\log n)\,\frac{\chi(n)}{n^s}.$$

Again, we rewrite this series using summation by parts as

$$\sum s_k \left[-k^{-s}\log k + (k+1)^{-s}\log(k+1)\right],$$

and an application of the mean-value theorem to the function $g(x) = x^{-s}\log x$ shows that the terms are $O(k^{-\delta/2-1})$, thus proving that the differentiated series converges uniformly for $s > \delta > 0$. Hence $L(s, \chi)$ is continuously differentiable for $s > 0$.

Now, observe that for all s large,

$$\begin{aligned}|L(s,\chi) - 1| &\leq 2q \sum_{n=2}^{\infty} n^{-s} \\ &\leq 2^{-s}O(1),\end{aligned}$$

and we can take $c = \log 2$, to see that $L(s, \chi) = 1 + O(e^{-cs})$ as $s \to \infty$. A similar argument also shows that $L'(s, \chi) = O(e^{-c's})$ as $s \to \infty$ with in fact $c' = c$, and the proof of the proposition is complete.

[2]For the formula of summation by parts, see Exercise 7 in Chapter 2.

With the facts gathered so far about $L(s, \chi)$ we are in a position to define the logarithm of the L-functions. This is done by integrating its logarithmic derivative. In other words, if χ is a non-trivial Dirichlet character and $s > 1$ we define[3]

$$\log_2 L(s, \chi) = -\int_s^\infty \frac{L'(t, \chi)}{L(t, \chi)}\, dt.$$

We know that $L(t, \chi) \neq 0$ for every $t > 1$ since it is given by a product (Proposition 3.2), and the integral is convergent because

$$\frac{L'(t, \chi)}{L(t, \chi)} = O(e^{-ct}),$$

which follows from the behavior at infinity of $L(t, \chi)$ and $L'(t, \chi)$ recorded earlier.

The following links the two logarithms.

Proposition 3.6 *If $s > 1$, then*

$$e^{\log_2 L(s,\chi)} = L(s, \chi).$$

Moreover

$$\log_2 L(s, \chi) = \sum_p \log_1 \left(\frac{1}{1 - \chi(p)/p^s} \right).$$

Proof. Differentiating $e^{-\log_2 L(s,\chi)} L(s, \chi)$ with respect to s gives

$$-\frac{L'(s, \chi)}{L(s, \chi)} e^{-\log_2 L(s,\chi)} L(s, \chi) + e^{-\log_2 L(s,\chi)} L'(s, \chi) = 0.$$

So $e^{-\log_2 L(s,\chi)} L(s, \chi)$ is constant, and this constant can be seen to be 1 by letting s tend to infinity. This proves the first conclusion.

To prove the equality between the logarithms, we fix s and take the exponential of both sides. The left-hand side becomes $e^{\log_2 L(s,\chi)} = L(s, \chi)$, and the right-hand side becomes

$$e^{\sum_p \log_1 \left(\frac{1}{1-\chi(p)/p^s} \right)} = \prod_p e^{\log_1 \left(\frac{1}{1-\chi(p)/p^s} \right)} = \prod_p \left(\frac{1}{1 - \chi(p)/p^s} \right) = L(s, \chi),$$

[3]The notation $\log_2$ used in this context should not be confused with the logarithm to the base 2.

by (i) in Proposition 3.1 and the Dirichlet product formula. Therefore, for each s there exists an integer $M(s)$ so that

$$\log_2 L(s,\chi) - \sum_p \log_1\left(\frac{1}{1-\chi(p)/p^s}\right) = 2\pi i M(s).$$

As the reader may verify, the left-hand side is continuous in s, and this implies the continuity of the function $M(s)$. But $M(s)$ is integer-valued so we conclude that $M(s)$ is constant, and this constant can be seen to be 0 by letting s go to infinity.

Putting together the work we have done so far gives rigorous meaning to the formal argument presented earlier. Indeed, the properties of $\log_1$ show that

$$\begin{aligned}\sum_p \log_1\left(\frac{1}{1-\chi(p)/p^s}\right) &= \sum_p \frac{\chi(p)}{p^s} + O\left(\sum_p \frac{1}{p^{2s}}\right)\\ &= \sum_p \frac{\chi(p)}{p^s} + O(1).\end{aligned}$$

Now if $L(1,\chi) \neq 0$ for a non-trivial Dirichlet character, then by its integral representation $\log_2 L(s,\chi)$ remains bounded as $s \to 1^+$. Thus the identity between the logarithms implies that $\sum_p \chi(p)p^{-s}$ remains bounded as $s \to 1^+$, which is the desired result. Therefore, to finish the proof of Dirichlet's theorem, we need to see that $L(1,\chi) \neq 0$ when χ is non-trivial.

3.3 Non-vanishing of the L-function

We now turn to a proof of the following deep result:

Theorem 3.7 *If $\chi \neq \chi_0$, then $L(1,\chi) \neq 0$.*

There are several proofs of this fact, some involving algebraic number theory (among them Dirichlet's original argument), and others involving complex analysis. Here we opt for a more elementary argument that requires no special knowledge of either of these areas. The proof splits in two cases, depending on whether χ is complex or real. A Dirichlet character is said to be **real** if it takes on only real values (that is, $+1$, -1, or 0) and **complex** otherwise. In other words, χ is real if and only if $\chi(n) = \overline{\chi(n)}$ for all integers n.

Case I: complex Dirichlet characters

This is the easier of the two cases. The proof is by contradiction, and we use two lemmas.

Lemma 3.8 *If* $s > 1$*, then*

$$\prod_{\chi} L(s, \chi) \geq 1,$$

where the product is taken over all Dirichlet characters. In particular the product is real-valued.

Proof. We have shown earlier that for $s > 1$

$$L(s, \chi) = \exp\left(\sum_p \log_1\left(\frac{1}{1 - \chi(p)p^{-s}}\right)\right).$$

Hence,

$$\begin{aligned}
\prod_{\chi} L(s, \chi) &= \exp\left(\sum_{\chi}\sum_p \log_1\left(\frac{1}{1 - \chi(p)p^{-s}}\right)\right) \\
&= \exp\left(\sum_{\chi}\sum_p \sum_{k=1}^{\infty} \frac{1}{k}\frac{\chi(p^k)}{p^{ks}}\right) \\
&= \exp\left(\sum_p \sum_{k=1}^{\infty}\sum_{\chi} \frac{1}{k}\frac{\chi(p^k)}{p^{ks}}\right).
\end{aligned}$$

Because of Lemma 2.2 (with $\ell = 1$) we have $\sum_{\chi} \chi(p^k) = \varphi(q)\delta_1(p^k)$, and hence

$$\prod_{\chi} L(s, \chi) = \exp\left(\varphi(q)\sum_p \sum_{k=1}^{\infty} \frac{1}{k}\frac{\delta_1(p^k)}{p^{ks}}\right) \geq 1,$$

since the term in the exponential is non-negative.

Lemma 3.9 *The following three properties hold:*

(i) *If* $L(1, \chi) = 0$*, then* $L(1, \overline{\chi}) = 0$.

(ii) *If* χ *is non-trivial and* $L(1, \chi) = 0$*, then*

$$|L(s, \chi)| \leq C|s - 1| \quad \textit{when } 1 \leq s \leq 2.$$

(iii) *For the trivial Dirichlet character* χ_0*, we have*

$$|L(s, \chi_0)| \leq \frac{C}{|s-1|} \quad \text{when } 1 < s \leq 2.$$

Proof. The first statement is immediate because $L(1, \overline{\chi}) = \overline{L(1, \chi)}$. The second statement follows from the mean-value theorem since $L(s, \chi)$ is continuously differentiable for $s > 0$ when χ is non-trivial. Finally, the last statement follows because by Proposition 3.3

$$L(s, \chi_0) = (1 - p_1^{-s})(1 - p_2^{-s}) \cdots (1 - p_N^{-s})\zeta(s),$$

and ζ satisfies the similar estimate (3).

We can now conclude the proof that $L(1, \chi) \neq 0$ for χ a non-trivial complex Dirichlet character. If not, say $L(1, \chi) = 0$, then we also have $L(1, \overline{\chi}) = 0$. Since $\chi \neq \overline{\chi}$, there are at least two terms in the product

$$\prod_{\chi} L(s, \chi),$$

that vanish like $|s-1|$ as $s \to 1^+$. Since only the trivial character contributes a term that grows, and this growth is no worse than $O(1/|s-1|)$, we find that the product goes to 0 as $s \to 1^+$, contradicting the fact that it is ≥ 1 by Lemma 3.8.

Case II: real Dirichlet characters

The proof that $L(1, \chi) \neq 0$ when χ is a non-trivial real Dirichlet character is very different from the earlier complex case. The method we shall exploit involves summation along hyperbolas. It is a curious fact that this method was introduced by Dirichlet himself, twelve years after the proof of his theorem on arithmetic progressions, to establish another famous result of his: the average order of the divisor function. However, he made no connection between the proofs of these two theorems. We will instead proceed by proving first Dirichlet's divisor theorem, as a simple example of the method of summation along hyperbolas. Then, we shall adapt these ideas to prove the fact that $L(1, \chi) \neq 0$. As a preliminary matter, we need to deal with some simple sums, and their corresponding integral analogues.

Sums vs. Integrals

Here we use the idea of comparing a sum with its corresponding integral, which already occurred in the estimate (3) for the zeta function.

Proposition 3.10 *If N is a positive integer, then:*

(i) $\displaystyle\sum_{1\leq n\leq N}\frac{1}{n}=\int_1^N\frac{dx}{x}+O(1)=\log N+O(1).$

(ii) *More precisely, there exists a real number γ, called Euler's constant, so that*

$$\sum_{1\leq n\leq N}\frac{1}{n}=\log N+\gamma+O(1/N).$$

Proof. It suffices to establish the more refined estimate given in part (ii). Let

$$\gamma_n=\frac{1}{n}-\int_n^{n+1}\frac{dx}{x}.$$

Since $1/x$ is decreasing, we clearly have

$$0\leq\gamma_n\leq\frac{1}{n}-\frac{1}{n+1}\leq\frac{1}{n^2},$$

so the series $\sum_{n=1}^{\infty}\gamma_n$ converges to a limit which we denote by γ. Moreover, if we estimate $\sum f(n)$ by $\int f(x)\,dx$, where $f(x)=1/x^2$, we find

$$\sum_{n=N+1}^{\infty}\gamma_n\leq\sum_{n=N+1}^{\infty}\frac{1}{n^2}\leq\int_N^{\infty}\frac{dx}{x^2}=O(1/N).$$

Therefore

$$\sum_{n=1}^{N}\frac{1}{n}-\int_1^N\frac{dx}{x}=\gamma-\sum_{n=N+1}^{\infty}\gamma_n+\int_N^{N+1}\frac{dx}{x},$$

and this last integral is $O(1/N)$ as $N\to\infty$.

Proposition 3.11 *If N is a positive integer, then*

$$\begin{aligned}\sum_{1\leq n\leq N}\frac{1}{n^{1/2}}&=\int_1^N\frac{dx}{x^{1/2}}+c'+O(1/N^{1/2})\\&=2N^{1/2}+c+O(1/N^{1/2}).\end{aligned}$$

The proof is essentially a repetition of the proof of the previous proposition, this time using the fact that

$$\left|\frac{1}{n^{1/2}}-\frac{1}{(n+1)^{1/2}}\right|\leq\frac{C}{n^{3/2}}.$$

This last inequality follows from the mean-value theorem applied to $f(x) = x^{-1/2}$, between $x = n$ and $x = n + 1$.

Hyperbolic sums

If F is a function defined on pairs of positive integers, there are three ways to calculate

$$S_N = \sum \sum F(m, n),$$

where the sum is taken over all pairs of positive integers (m, n) which satisfy $mn \leq N$.

We may carry out the summation in any one of the following three ways. (See Figure 2.)

(a) Along hyperbolas:

$$S_N = \sum_{1 \leq k \leq N} \left(\sum_{nm=k} F(m, n) \right)$$

(b) Vertically:

$$S_N = \sum_{1 \leq m \leq N} \left(\sum_{1 \leq n \leq N/m} F(m, n) \right)$$

(c) Horizontally:

$$S_N = \sum_{1 \leq n \leq N} \left(\sum_{1 \leq m \leq N/n} F(m, n) \right)$$

It is a remarkable fact that one can obtain interesting conclusions from the obvious fact that these three methods of summation give the same sum. We apply this idea first in the study of the divisor problem.

Intermezzo: the divisor problem

For a positive integer k, let $d(k)$ denote the number of positive divisors of k. For example,

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
$d(k)$	1	2	2	3	2	4	2	4	3	4	2	6	2	4	4	5	2

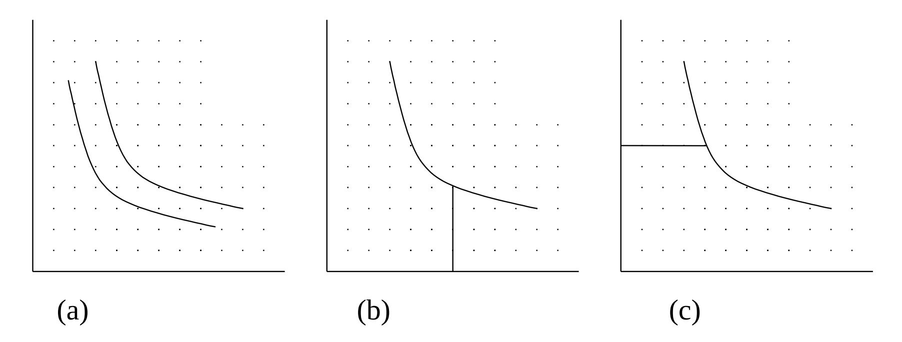


Figure 2. The three methods of summation

One observes that the behavior of $d(k)$ as k tends to infinity is rather irregular, and in fact, it does not seem possible to approximate $d(k)$ by a simple analytic expression in k. However, it is natural to inquire about the average size of $d(k)$. In other words, one might ask, what is the behavior of

$$\frac{1}{N}\sum_{k=1}^{N} d(k) \qquad \text{as } N \to \infty?$$

The answer was provided by Dirichlet, who made use of hyperbolic sums. Indeed, we observe that

$$d(k) = \sum_{nm=k,\ 1\leq n,m} 1.$$

Theorem 3.12 *If k is a positive integer, then*

$$\frac{1}{N}\sum_{k=1}^{N} d(k) = \log N + O(1).$$

More precisely,

$$\frac{1}{N}\sum_{k=1}^{N} d(k) = \log N + (2\gamma - 1) + O(1/N^{1/2}),$$

where γ is Euler's constant.

Proof. Let $S_N = \sum_{k=1}^{N} d(k)$. We observed that summing $F = 1$ along hyperbolas gives S_N. Summing vertically, we find

$$S_N = \sum_{1\leq m\leq N}\ \sum_{1\leq n\leq N/m} 1.$$

But $\sum_{1\leq n\leq N/m} 1 = [N/m] = N/m + O(1)$, where $[x]$ denote the greatest integer $\leq x$. Therefore

$$S_N = \sum_{1\leq m\leq N} (N/m + O(1)) = N \left(\sum_{1\leq m\leq N} 1/m \right) + O(N).$$

Hence, by part (i) of Proposition 3.10,

$$\frac{S_N}{N} = \log N + O(1)$$

which gives the first conclusion.

For the more refined estimate we proceed as follows. Consider the three regions I, II, and III shown in Figure 3. These are defined by

$$\begin{aligned} I &= \{1 \leq m < N^{1/2},\ N^{1/2} < n \leq N/m\}, \\ II &= \{1 \leq m \leq N^{1/2},\ 1 \leq n \leq N^{1/2}\}, \\ III &= \{N^{1/2} < m \leq N/n,\ 1 \leq n < N^{1/2}\}. \end{aligned}$$

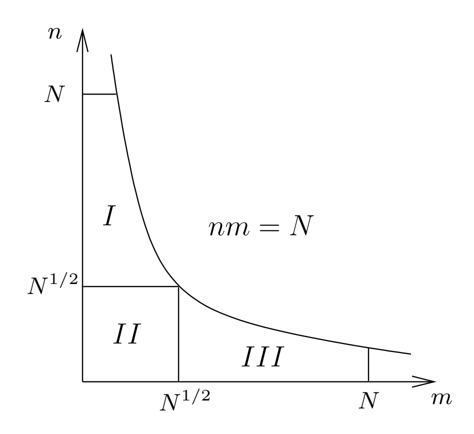


Figure 3. The three regions I, II, and III

If S_I, S_{II}, and S_{III} denote the sums taken over the regions I, II, and III, respectively, then

$$\begin{aligned} S_N &= S_I + S_{II} + S_{III} \\ &= 2(S_I + S_{II}) - S_{II}, \end{aligned}$$

since by symmetry $S_I = S_{III}$. Now we sum vertically, and use (ii) of Proposition 3.10 to obtain

$$\begin{aligned} S_I + S_{II} &= \sum_{1\le m\le N^{1/2}} \left(\sum_{1\le n\le N/m} 1 \right) \\ &= \sum_{1\le m\le N^{1/2}} [N/m] \\ &= \sum_{1\le m\le N^{1/2}} (N/m + O(1)) \\ &= N \left(\sum_{1\le m\le N^{1/2}} 1/m \right) + O(N^{1/2}) \\ &= N\log N^{1/2} + N\gamma + O(N^{1/2}). \end{aligned}$$

Finally, S_{II} corresponds to a square so

$$S_{II} = \sum_{1\le m\le N^{1/2}} \sum_{1\le n\le N^{1/2}} 1 = [N^{1/2}]^2 = N + O(N^{1/2}).$$

Putting these estimates together and dividing by N yields the more refined statement in the theorem.

Non-vanishing of the L-function

Our essential application of summation along hyperbolas is to the main point of this section, namely that $L(1,\chi) \neq 0$ for a non-trivial real Dirichlet character χ.

Given such a character, let

$$F(m,n) = \frac{\chi(n)}{(nm)^{1/2}},$$

and define

$$S_N = \sum\sum F(m,n),$$

where the sum is over all integers $m, n \ge 1$ that satisfy $mn \le N$.

Proposition 3.13 *The following statements are true:*

(i) $S_N \ge c\log N$ *for some constant* $c > 0$.

(ii) $S_N = 2N^{1/2}L(1,\chi) + O(1)$.

It suffices to prove the proposition, since the assumption $L(1, \chi) = 0$ would give an immediate contradiction.

We first sum along hyperbolas. Observe that

$$\sum_{nm=k} \frac{\chi(n)}{(nm)^{1/2}} = \frac{1}{k^{1/2}} \sum_{n|k} \chi(n).$$

For conclusion (i) it will be enough to show the following lemma.

Lemma 3.14 $\displaystyle\sum_{n|k} \chi(n) \geq \begin{cases} 0 & \textit{for all } k \\ 1 & \textit{if } k = \ell^2 \textit{ for some } \ell \in \mathbb{Z}. \end{cases}$

From the lemma, we then get

$$S_N \geq \sum_{k=\ell^2,\ \ell \leq N^{1/2}} \frac{1}{k^{1/2}} \geq c \log N,$$

where the last inequality follows from (i) in Proposition 3.10.

The proof of the lemma is simple. If k is a power of a prime, say $k = p^a$, then the divisors of k are $1, p, p^2, \ldots, p^a$ and

$$\begin{aligned} \sum_{n|k} \chi(n) &= \chi(1) + \chi(p) + \chi(p^2) + \cdots + \chi(p^a) \\ &= 1 + \chi(p) + \chi(p)^2 + \cdots + \chi(p)^a. \end{aligned}$$

So this sum is equal to

$$\begin{cases} a+1 & \text{if } \chi(p) = 1, \\ 1 & \text{if } \chi(p) = -1 \text{ and } a \text{ is even}, \\ 0 & \text{if } \chi(p) = -1 \text{ and } a \text{ is odd}, \\ 1 & \text{if } \chi(p) = 0, \text{ that is } p|q. \end{cases}$$

In general, if $k = p_1^{a_1} \cdots p_N^{a_N}$, then any divisor of k is of the form $p_1^{b_1} \cdots p_N^{b_N}$ where $0 \leq b_j \leq a_j$ for all j. Therefore, the multiplicative property of χ gives

$$\sum_{n|k} \chi(n) = \prod_{j=1}^{N} \left(\chi(1) + \chi(p_j) + \chi(p_j^2) + \cdots + \chi(p_j^{a_j}) \right),$$

and the proof is complete.

To prove the second statement in the proposition, we write

$$S_N = S_I + (S_{II} + S_{III}),$$

where the sums S_I, S_{II}, and S_{III} were defined earlier (see also Figure 3). We evaluate S_I by summing vertically, and $S_{II} + S_{III}$ by summing horizontally. In order to carry this out we need the following simple results.

Lemma 3.15 *For all integers $0 < a < b$ we have*

(i) $\displaystyle\sum_{n=a}^{b} \frac{\chi(n)}{n^{1/2}} = O(a^{-1/2})$,

(ii) $\displaystyle\sum_{n=a}^{b} \frac{\chi(n)}{n} = O(a^{-1})$.

Proof. This argument is similar to the proof of Proposition 3.4; we use summation by parts. Let $s_n = \sum_{1 \leq k \leq n} \chi(k)$, and remember that $|s_n| \leq q$ for all n. Then

$$\begin{aligned}\sum_{n=a}^{b} \frac{\chi(n)}{n^{1/2}} &= \sum_{n=a}^{b-1} s_n \left[n^{-1/2} - (n+1)^{-1/2}\right] + O(a^{-1/2}) \\ &= O\left(\sum_{n=a}^{\infty} n^{-3/2}\right) + O(a^{-1/2}).\end{aligned}$$

By comparing the sum $\sum_{n=a}^{\infty} n^{-3/2}$ with the integral of $f(x) = x^{-3/2}$, we find that the former is also $O(a^{-1/2})$.

A similar argument establishes (ii).

We may now finish the proof of the proposition. Summing vertically we find

$$S_I = \sum_{m < N^{1/2}} \frac{1}{m^{1/2}} \left(\sum_{N^{1/2} < n \leq N/m} \chi(n)/n^{1/2} \right).$$

The lemma together with Proposition 3.11 shows that $S_I = O(1)$. Finally

we sum horizontally to get

$$\begin{aligned}
S_{II} + S_{III} &= \sum_{1 \leq n \leq N^{1/2}} \frac{\chi(n)}{n^{1/2}} \left(\sum_{m \leq N/n} 1/m^{1/2} \right) \\
&= \sum_{1 \leq n \leq N^{1/2}} \frac{\chi(n)}{n^{1/2}} \left\{ 2(N/n)^{1/2} + c + O((n/N)^{1/2}) \right\} \\
&= 2N^{1/2} \sum_{1 \leq n \leq N^{1/2}} \frac{\chi(n)}{n} + c \sum_{1 \leq n \leq N^{1/2}} \frac{\chi(n)}{n^{1/2}} \\
&\qquad + O\left(\frac{1}{N^{1/2}} \sum_{1 \leq n \leq N^{1/2}} 1 \right) \\
&= A + B + C.
\end{aligned}$$

Now observe that the lemma, together with the definition of $L(s, \chi)$, implies

$$A = 2N^{1/2} L(1, \chi) + O(N^{1/2} N^{-1/2}).$$

Moreover, part (i) of the lemma gives $B = O(1)$, and we also clearly have $C = O(1)$. Thus $S_N = 2N^{1/2} L(1, \chi) + O(1)$, which is part (ii) in Proposition 3.13.

This completes the proof that $L(1, \chi) \neq 0$, and thus the proof of Dirichlet's theorem.

4 Exercises

1. Prove that there are infinitely many primes by observing that if there were only finitely many, $p_1, \ldots, p_N$, then

$$\prod_{j=1}^{N} \frac{1}{1 - 1/p_j} \geq \sum_{n=1}^{\infty} \frac{1}{n}.$$

2. In the text we showed that there are infinitely many primes of the form $4k + 3$ by a modification of Euclid's original argument. Adapt this technique to prove the similar result for primes of the form $3k + 2$, and for those of the form $6k + 5$.

3. Prove that if p and q are relatively prime, then $\mathbb{Z}^*(p) \times \mathbb{Z}^*(q)$ is isomorphic to $\mathbb{Z}^*(pq)$.

4. Let $\varphi(n)$ denote the number of positive integers $\leq n$ that are relatively prime to n. Use the previous exercise to show that if n and m are relatively prime, then

$$\varphi(nm) = \varphi(n)\varphi(m).$$

One can give a formula for the Euler phi-function as follows:

(a) Calculate $\varphi(p)$ when p is prime by counting the number of elements in $\mathbb{Z}^*(p)$.

(b) Give a formula for $\varphi(p^k)$ when p is prime and $k \geq 1$ by counting the number of elements in $\mathbb{Z}^*(p^k)$.

(c) Show that

$$\varphi(n) = n \prod_i \left(1 - \frac{1}{p_i}\right),$$

where p_i are the primes that divide n.

5. If n is a positive integer, show that

$$n = \sum_{d|n} \varphi(d),$$

where φ is the Euler phi-function.
[Hint: There are precisely $\varphi(n/d)$ integers $1 \leq m \leq n$ with $\gcd(m, n) = d$.]

6. Write down the characters of the groups $\mathbb{Z}^*(3)$, $\mathbb{Z}^*(4)$, $\mathbb{Z}^*(5)$, $\mathbb{Z}^*(6)$, and $\mathbb{Z}^*(8)$.

(a) Which ones are real, or complex?

(b) Which ones are even, or odd? (A character is even if $\chi(-1) = 1$, and odd otherwise).

7. Recall that for $|z| < 1$,

$$\log_1 \left(\frac{1}{1-z}\right) = \sum_{k \geq 1} \frac{z^k}{k}.$$

We have seen that

$$e^{\log_1\left(\frac{1}{1-z}\right)} = \frac{1}{1-z}.$$

(a) Show that if $w = 1/(1-z)$, then $|z| < 1$ if and only if $\text{Re}(w) > 1/2$.

(b) Show that if $\text{Re}(w) > 1/2$ and $w = \rho e^{i\varphi}$ with $\rho > 0$, $|\varphi| < \pi$, then

$$\log_1 w = \log \rho + i\varphi.$$

[Hint: If $e^{\zeta} = w$, then the real part of ζ is uniquely determined and its imaginary part is determined modulo 2π.]

8. Let ζ denote the zeta function defined for $s > 1$.

(a) Compare $\zeta(s)$ with $\int_1^{\infty} x^{-s}\,dx$ to show that

$$\zeta(s) = \frac{1}{s-1} + O(1) \qquad \text{as } s \to 1^+.$$

(b) Prove as a consequence that

$$\sum_p \frac{1}{p^s} = \log\left(\frac{1}{s-1}\right) + O(1) \qquad \text{as } s \to 1^+.$$

9. Let χ_0 denote the trivial Dirichlet character mod q, and $p_1, \ldots, p_k$ the distinct prime divisors of q. Recall that $L(s, \chi_0) = (1 - p_1^{-s}) \cdots (1 - p_k^{-s})\zeta(s)$, and show as a consequence

$$L(s, \chi_0) = \frac{\varphi(q)}{q} \frac{1}{s-1} + O(1) \qquad \text{as } s \to 1^+.$$

[Hint: Use the asymptotics for ζ in Exercise 8.]

10. Show that if ℓ is relatively prime to q, then

$$\sum_{p \equiv \ell} \frac{1}{p^s} = \frac{1}{\varphi(q)} \log\left(\frac{1}{s-1}\right) + O(1) \qquad \text{as } s \to 1^+.$$

This is a quantitative version of Dirichlet's theorem.

[Hint: Recall (4).]

11. Use the characters for $\mathbb{Z}^*(3)$, $\mathbb{Z}^*(4)$, $\mathbb{Z}^*(5)$, and $\mathbb{Z}^*(6)$ to verify directly that $L(1, \chi) \neq 0$ for all non-trivial Dirichlet characters modulo q when $q = 3, 4, 5$, and 6.

[Hint: Consider in each case the appropriate alternating series.]

12. Suppose χ is real and non-trivial; assuming the theorem that $L(1, \chi) \neq 0$, show directly that $L(1, \chi) > 0$.

[Hint: Use the product formula for $L(s, \chi)$.]

13. Let $\{a_n\}_{n=-\infty}^{\infty}$ be a sequence of complex numbers such that $a_n = a_m$ if $n = m \bmod q$. Show that the series

$$\sum_{n=1}^{\infty} \frac{a_n}{n}$$

converges if and only if $\sum_{n=1}^{q} a_n = 0$.

[Hint: Sum by parts.]

14. The series

$$F(\theta) = \sum_{|n| \neq 0} \frac{e^{in\theta}}{n}, \qquad \text{for } |\theta| < \pi,$$

converges for every θ and is the Fourier series of the function defined on $[-\pi, \pi]$ by $F(0) = 0$ and

$$F(\theta) = \begin{cases} i(-\pi - \theta) & \text{if } -\pi \leq \theta < 0 \\ i(\pi - \theta) & \text{if } 0 < \theta \leq \pi, \end{cases}$$

and extended by periodicity (period 2π) to all of $\mathbb{R}$ (see Exercise 8 in Chapter 2).

Show also that if $\theta \neq 0 \bmod 2\pi$, then the series

$$E(\theta) = \sum_{n=1}^{\infty} \frac{e^{in\theta}}{n}$$

converges, and that

$$E(\theta) = \frac{1}{2} \log \left(\frac{1}{2 - 2\cos\theta} \right) + \frac{i}{2} F(\theta).$$

15. To sum the series $\sum_{n=1}^{\infty} a_n/n$, with $a_n = a_m$ if $n = m \bmod q$ and $\sum_{n=1}^{q} a_n = 0$, proceed as follows.

(a) Define

$$A(m) = \sum_{n=1}^{q} a_n \zeta^{-mn} \qquad \text{where } \zeta = e^{2\pi i/q}.$$

Note that $A(q) = 0$. With the notation of the previous exercise, prove that

$$\sum_{n=1}^{\infty} \frac{a_n}{n} = \frac{1}{q} \sum_{m=1}^{q-1} A(m) E(2\pi m/q).$$

[Hint: Use Fourier inversion on $\mathbb{Z}(q)$.]

(b) If $\{a_m\}$ is odd, $(a_{-m} = -a_m)$ for $m \in \mathbb{Z}$, observe that $a_0 = a_q = 0$ and show that

$$A(m) = \sum_{1 \leq n < q/2} a_n (\zeta^{-mn} - \zeta^{mn}).$$

(c) Still assuming that $\{a_m\}$ is odd, show that

$$\sum_{n=1}^{\infty} \frac{a_n}{n} = \frac{1}{2q} \sum_{m=1}^{q-1} A(m) F(2\pi m/q).$$

[Hint: Define $\tilde{A}(m) = \sum_{n=1}^{q} a_n \zeta^{mn}$ and apply the Fourier inversion formula.]

16. Use the previous exercises to show that

$$\frac{\pi}{3\sqrt{3}} = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{5} + \frac{1}{7} - \frac{1}{8} + \cdots,$$

which is $L(1, \chi)$ for the non-trivial (odd) Dirichlet character modulo 3.

5 Problems

1.* Here are other series that can be summed by the methods in Exercise 15.

(a) For the non-trivial Dirichlet character modulo 6, $L(1, \chi)$ equals

$$\frac{\pi}{2\sqrt{3}} = 1 - \frac{1}{5} + \frac{1}{7} - \frac{1}{11} + \frac{1}{13} + \cdots.$$

(b) If χ is the odd Dirichlet character modulo 8, then $L(1, \chi)$ equals

$$\frac{\pi}{2\sqrt{2}} = 1 + \frac{1}{3} - \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \frac{1}{11} \cdots.$$

(c) For an odd Dirichlet character modulo 7, $L(1, \chi)$ equals

$$\frac{\pi}{\sqrt{7}} = 1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} - \frac{1}{6} \cdots.$$

(d) For an even Dirichlet character modulo 8, $L(1, \chi)$ equals

$$\frac{\log(1+\sqrt{2})}{\sqrt{2}} = 1 - \frac{1}{3} - \frac{1}{5} + \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \cdots .$$

(e) For an even Dirichlet character modulo 5, $L(1, \chi)$ equals

$$\frac{2}{\sqrt{5}} \log\left(\frac{1+\sqrt{5}}{2}\right) = 1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{6} - \frac{1}{7} - \frac{1}{8} + \frac{1}{9} + \frac{1}{11} \cdots .$$

2. Let $d(k)$ denote the number of positive divisors of k.

(a) Show that if $k = p_1^{a_1} \cdots p_n^{a_n}$ is the prime factorization of k, then

$$d(k) = (a_1 + 1) \cdots (a_n + 1).$$

Although Theorem 3.12 shows that on "average" $d(k)$ is of the order of $\log k$, prove the following on the basis of (a):

(b) $d(k) = 2$ for infinitely many k.

(c) For any positive integer N, there is a constant $c > 0$ so that $d(k) \geq c(\log k)^N$ for infinitely many k. [Hint: Let $p_1, \ldots, p_N$ be N distinct primes, and consider k of the form $(p_1 p_2 \cdots p_N)^m$ for $m = 1, 2, \ldots$.]

3. Show that if p is relatively prime to q, then

$$\prod_{\chi} \left(1 - \frac{\chi(p)}{p^s}\right) = \left(\frac{1}{1 - p^{fs}}\right)^g,$$

where $g = \varphi(q)/f$, and f is the order of p in $\mathbb{Z}^*(q)$ (that is, the smallest n for which $p^n \equiv 1 \mod q$). Here the product is taken over all Dirichlet characters modulo q.

4. Prove as a consequence of the previous problem that

$$\prod_{\chi} L(s, \chi) = \sum_{n \geq 1} \frac{a_n}{n^s},$$

where $a_n \geq 0$, and the product is over all Dirichlet characters modulo q.

Appendix : Integration

This appendix is meant as a quick review of the definition and main properties of the Riemann integral on $\mathbb{R}$, and integration of appropriate continuous functions on $\mathbb{R}^d$. Our exposition is brief since we assume that the reader already has some familiarity with this material.

We begin with the theory of Riemann integration on a closed and bounded interval on the real line. Besides the standard results about the integral, we also discuss the notion of sets of measure 0, and give a necessary and sufficient condition on the set of discontinuities of a function that guarantee its integrability.

We also discuss multiple and repeated integrals. In particular, we extend the notion of integration to the entire space $\mathbb{R}^d$ by restricting ourselves to functions that decay fast enough at infinity.

1 Definition of the Riemann integral

Let f be a *bounded* real-valued function defined on the closed interval $[a, b] \subset \mathbb{R}$. By a **partition** P of $[a, b]$ we mean a finite sequence of numbers $x_0, x_1, \ldots, x_N$ with

$$a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b.$$

Given such a partition, we let I_j denote the interval $[x_{j-1}, x_j]$ and write $|I_j|$ for its length, namely $|I_j| = x_j - x_{j-1}$. We define the upper and lower sums of f with respect to P by

$$\mathcal{U}(P, f) = \sum_{j=1}^{N} [\sup_{x \in I_j} f(x)]\, |I_j| \quad \text{and} \quad \mathcal{L}(P, f) = \sum_{j=1}^{N} [\inf_{x \in I_j} f(x)]\, |I_j|.$$

Note that the infimum and supremum exist because by assumption, f is bounded. Clearly $\mathcal{U}(P, f) \geq \mathcal{L}(P, f)$, and the function f is said to be **Riemann integrable**, or simply **integrable**, if for every $\epsilon > 0$ there exists a partition P such that

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon.$$

To define the value of the integral of f, we need to make a simple yet important observation. A partition P' is said to be a **refinement** of the partition P if P' is obtained from P by adding points. Then, adding one

point at a time, it is easy to check that

$$\mathcal{U}(P', f) \leq \mathcal{U}(P, f) \quad \text{and} \quad \mathcal{L}(P', f) \geq \mathcal{L}(P, f).$$

From this, we see that if P_1 and P_2 are two partitions of $[a, b]$, then

$$\mathcal{U}(P_1, f) \geq \mathcal{L}(P_2, f),$$

since it is possible to take P' as a common refinement of both P_1 and P_2 to obtain

$$\mathcal{U}(P_1, f) \geq \mathcal{U}(P', f) \geq \mathcal{L}(P', f) \geq \mathcal{L}(P_2, f).$$

Since f is bounded we see that both

$$U = \inf_P \mathcal{U}(P, f) \quad \text{and} \quad L = \sup_P \mathcal{L}(P, f)$$

exist (where the infimum and supremum are taken over all partitions of $[a, b]$), and also that $U \geq L$. Moreover, if f is integrable we must have $U = L$, and we *define* $\int_a^b f(x)\,dx$ to be this common value.

Finally, a bounded complex-valued function $f = u + iv$ is said to be integrable if its real and imaginary parts u and v are integrable, and we define

$$\int_a^b f(x)\,dx = \int_a^b u(x)\,dx + i \int_a^b v(x)\,dx.$$

For example, the constants are integrable functions and it is clear that if $c \in \mathbb{C}$, then $\int_a^b c\,dx = c(b-a)$. Also, continuous functions are integrable. This is because a continuous function on a closed and bounded interval $[a, b]$ is uniformly continuous, that is, given $\epsilon > 0$ there exists δ such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \epsilon$. So if we choose n with $(b - a)/n < \delta$, then the partition P given by

$$a,\ a + \frac{b-a}{n}, \ldots,\ a + k\frac{b-a}{n}, \ldots,\ a + (n-1)\frac{b-a}{n},\ b$$

satisfies $\mathcal{U}(P, f) - \mathcal{L}(P, f) \leq \epsilon(b - a)$.

1.1 Basic properties

Proposition 1.1 *If f and g are integrable on $[a, b]$, then:*

(i) *$f + g$ is integrable, and $\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$.*

(ii) *If* $c \in \mathbb{C}$, *then* $\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$.

(iii) *If* f *and* g *are real-valued and* $f(x) \leq g(x)$, *then* $\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$.

(iv) *If* $c \in [a,b]$, *then* $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$.

Proof. For property (i) we may assume that f and g are real-valued. If P is a partition of $[a,b]$, then

$$\mathcal{U}(P,f+g) \leq \mathcal{U}(P,f) + \mathcal{U}(P,g) \quad \text{and} \quad \mathcal{L}(P,f+g) \geq \mathcal{L}(P,f) + \mathcal{L}(P,g).$$

Given $\epsilon > 0$, there exist partitions P_1 and P_2 such that $\mathcal{U}(P_1,f) - \mathcal{L}(P_1,f) < \epsilon$ and $\mathcal{U}(P_2,g) - \mathcal{L}(P_2,g) < \epsilon$, so that if P_0 is a common refinement of P_1 and P_2, we get

$$\mathcal{U}(P_0,f+g) - \mathcal{L}(P_0,f+g) < 2\epsilon.$$

So $f+g$ is integrable, and if we let $I = \inf_P \mathcal{U}(P,f+g) = \sup_P \mathcal{L}(P,f+g)$, then we see that

$$\begin{aligned} I &\leq \mathcal{U}(P_0,f+g) + 2\epsilon \leq \mathcal{U}(P_0,f) + \mathcal{U}(P_0,g) + 2\epsilon \\ &\leq \int_a^b f(x)\,dx + \int_a^b g(x)\,dx + 4\epsilon. \end{aligned}$$

Similarly $I \geq \int_a^b f(x)\,dx + \int_a^b g(x)\,dx - 4\epsilon$, which proves that $\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$. The second and third parts of the proposition are just as easy to prove. For the last property, simply refine partitions of $[a,b]$ by adding the point c.

Another important property we need to prove is that fg is integrable whenever f and g are integrable.

Lemma 1.2 *If* f *is real-valued integrable on* $[a,b]$ *and* φ *is a real-valued continuous function on* $\mathbb{R}$, *then* $\varphi \circ f$ *is also integrable on* $[a,b]$.

Proof. Let $\epsilon > 0$ and remember that f is bounded, say $|f| \leq M$. Since φ is uniformly continuous on $[-M,M]$ we may choose $\delta > 0$ so that if $s,t \in [-M,M]$ and $|s-t| < \delta$, then $|\varphi(s) - \varphi(t)| < \epsilon$. Now choose a partition $P = \{x_0, \ldots, x_N\}$ of $[a,b]$ with $\mathcal{U}(P,f) - \mathcal{L}(P,f) < \delta^2$. Let $I_j = [x_{j-1}, x_j]$ and distinguish two classes: we write $j \in \Lambda$ if $\sup_{x\in I_j} f(x) - \inf_{x \in I_j} f(x) < \delta$ so that by construction

$$\sup_{x\in I_j} \varphi \circ f(x) - \inf_{x\in I_j} \varphi \circ f(x) < \epsilon.$$

Otherwise, we write $j \in \Lambda'$ and note that

$$\delta \sum_{j\in\Lambda'} |I_j| \leq \sum_{j\in\Lambda'} [\sup_{x\in I_j} f(x) - \inf_{x\in I_j} f(x)]\, |I_j| \leq \delta^2$$

so $\sum_{j\in\Lambda'} |I_j| < \delta$. Therefore, separating the cases $j \in \Lambda$ and $j \in \Lambda'$ we find that

$$\mathcal{U}(P, \varphi \circ f) - \mathcal{L}(P, \varphi \circ f) \leq \epsilon(b-a) + 2\mathcal{B}\delta,$$

where $\mathcal{B}$ is a bound for φ on $[-M, M]$. Since we can also choose $\delta < \epsilon$, we see that the proposition is proved.

¿From the lemma we get the following facts:

- If f and g are integrable on $[a, b]$, then the product fg is integrable on $[a, b]$.

This follows from the lemma with $\varphi(t) = t^2$, and the fact that $fg = \frac{1}{4}\left([f+g]^2 - [f-g]^2\right)$.

- If f is integrable on $[a, b]$, then the function $|f|$ is integrable, and $\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$.

We can take $\varphi(t) = |t|$ to see that $|f|$ is integrable. Moreover, the inequality follows from (iii) in Proposition 1.1.

We record two results that imply integrability.

Proposition 1.3 *A bounded monotonic function f on an interval $[a, b]$ is integrable.*

Proof. We may assume without loss of generality that $a = 0$, $b = 1$, and f is monotonically increasing. Then, for each N, we choose the uniform partition P_N given by $x_j = j/N$ for all $j = 0, \ldots, N$. If $\alpha_j = f(x_j)$, then we have

$$\mathcal{U}(P_N, f) = \frac{1}{N}\sum_{j=1}^{N} \alpha_j \quad \text{and} \quad \mathcal{L}(P_N, f) = \frac{1}{N}\sum_{j=1}^{N} \alpha_{j-1}.$$

Therefore, if $|f(x)| \leq B$ for all x we have

$$\mathcal{U}(P_N, f) - \mathcal{L}(P_N, f) = \frac{\alpha_N - \alpha_0}{N} \leq \frac{2B}{N},$$

and the proposition is proved.

Proposition 1.4 *Let f be a bounded function on the compact interval $[a, b]$. If $c \in (a, b)$, and if for all small $\delta > 0$ the function f is integrable on the intervals $[a, c - \delta]$ and $[c + \delta, b]$, then f is integrable on $[a, b]$.*

Proof. Suppose $|f| \leq M$ and let $\epsilon > 0$. Choose $\delta > 0$ (small) so that $4\delta M \leq \epsilon/3$. Now let P_1 and P_2 be partitions of $[a, c - \delta]$ and $[c + \delta, b]$ so that for each $i = 1, 2$ we have $\mathcal{U}(P_i, f) - \mathcal{L}(P_i, f) < \epsilon/3$. This is possible since f is integrable on each one of the intervals. Then by taking as a partition $P = P_1 \cup \{c - \delta\} \cup \{c + \delta\} \cup P_2$ we immediately see that $\mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon$.

We end this section with a useful approximation lemma. Recall that a function on the circle is the same as a 2π-periodic function on $\mathbb{R}$.

Lemma 1.5 *Suppose f is integrable on the circle and f is bounded by B. Then there exists a sequence $\{f_k\}_{k=1}^{\infty}$ of continuous functions on the circle so that*

$$\sup_{x \in [-\pi, \pi]} |f_k(x)| \leq B \quad \textit{for all } k = 1, 2, \ldots,$$

and

$$\int_{-\pi}^{\pi} |f(x) - f_k(x)|\, dx \to 0 \quad \textit{as } k \to \infty.$$

Proof. Assume f is real-valued (in general apply the following argument to the real and imaginary parts separately). Given $\epsilon > 0$, we may choose a partition $-\pi = x_0 < x_1 < \cdots < x_N = \pi$ of the interval $[-\pi, \pi]$ so that the upper and lower sums of f differ by at most ϵ. Denote by f^* the step function defined by

$$f^*(x) = \sup_{x_{j-1} \leq y \leq x_j} f(y) \quad \text{if } x \in [x_{j-1}, x_j) \text{ for } 1 \leq j \leq N.$$

By construction we have $|f^*| \leq B$, and moreover

$$(1) \qquad \int_{-\pi}^{\pi} |f^*(x) - f(x)|\, dx = \int_{-\pi}^{\pi} (f^*(x) - f(x))\, dx < \epsilon.$$

Now we can modify f^* to make it continuous and periodic yet still approximate f in the sense of the lemma. For small $\delta > 0$, let $\tilde{f}(x) = f^*(x)$ when the distance of x from any of the division points $x_0, \ldots, x_N$ is $\geq \delta$. In the δ-neighborhood of x_j for $j = 1, \ldots, N - 1$, define $\tilde{f}(x)$ to be the linear function for which $\tilde{f}(x_j \pm \delta) = f^*(x_j \pm \delta)$. Near $x_0 = -\pi$, $\tilde{f}$

Ibookroot October 20, 2007

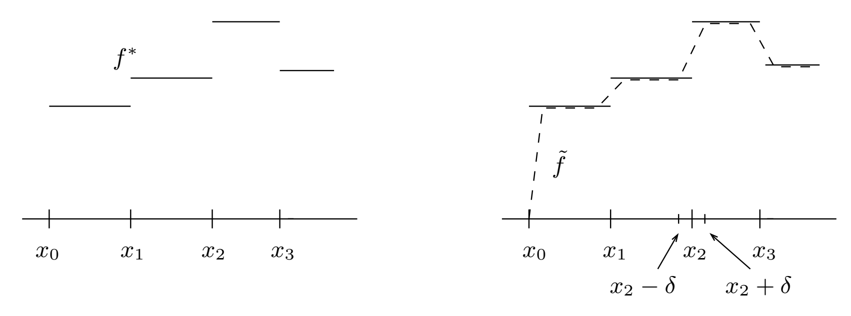


Figure 1. Portions of the functions f^* and $\tilde{f}$

is linear with $\tilde{f}(-\pi) = 0$ and $\tilde{f}(-\pi+\delta) = f^*(-\pi+\delta)$. Similarly, near $x_N = \pi$ the function $\tilde{f}$ is linear with $\tilde{f}(\pi) = 0$ and $\tilde{f}(\pi-\delta) = f^*(\pi-\delta)$. In Figure 1 we illustrate the situation near $x_0 = -\pi$. In the second picture the graph of $\tilde{f}$ is shifted slightly below to clarify the situation.

Then, since $\tilde{f}(-\pi) = \tilde{f}(\pi)$, we may extend $\tilde{f}$ to a continuous and 2π-periodic function on $\mathbb{R}$. The absolute value of this extension is also bounded by B. Moreover, $\tilde{f}$ differs from f^* only in the N intervals of length 2δ surrounding the division points. Thus

$$\int_{-\pi}^{\pi} |f^*(x) - \tilde{f}(x)|\, dx \leq 2BN(2\delta).$$

If we choose δ sufficiently small, we get

$$\int_{-\pi}^{\pi} |f^*(x) - \tilde{f}(x)|\, dx < \epsilon. \tag{2}$$

As a result, equations (1), (2), and the triangle inequality yield

$$\int_{-\pi}^{\pi} |f(x) - \tilde{f}(x)|\, dx < 2\epsilon.$$

Denoting by f_k the $\tilde{f}$ so constructed, when $2\epsilon = 1/k$, we see that the sequence $\{f_k\}$ has the properties required by the lemma.

1.2 Sets of measure zero and discontinuities of integrable functions

We observed that all continuous functions are integrable. By modifying the argument slightly, one can show that all piecewise continuous functions are also integrable. In fact, this is a consequence of Proposition 1.4

applied finitely many times. We now turn to a more careful study of the discontinuities of integrable functions.

We start with a definition[1]: a subset E of $\mathbb{R}$ is said to have **measure** 0 if for every $\epsilon > 0$ there exists a countable family of open intervals $\{I_k\}_{k=1}^{\infty}$ such that

(i) $E \subset \bigcup_{k=1}^{\infty} I_k$,

(ii) $\sum_{k=1}^{\infty} |I_k| < \epsilon$, where $|I_k|$ denotes the length of the interval I_k.

The first condition says that the union of the intervals covers E, and the second that this union is small. The reader will have no difficulty proving that any finite set of points has measure 0. A more subtle argument is needed to prove that a countable set of points has measure 0. In fact, this result is contained in the following lemma.

Lemma 1.6 *The union of countably many sets of measure* 0 *has measure* 0.

Proof. Say $E_1, E_2, \ldots$ are sets of measure 0, and let $E = \cup_{i=1}^{\infty} E_i$. Let $\epsilon > 0$, and for each i choose open interval $I_{i,1}, I_{i,2}, \ldots$ so that

$$E_i \subset \bigcup_{k=1}^{\infty} I_{i,k} \quad \text{and} \quad \sum_{k=1}^{\infty} |I_{i,k}| < \epsilon/2^i.$$

Now clearly we have $E \subset \bigcup_{i,k=1}^{\infty} I_{i,k}$, and

$$\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} |I_{i,k}| \leq \sum_{i=1}^{\infty} \frac{\epsilon}{2^i} \leq \epsilon,$$

as was to be shown.

An important observation is that if E has measure 0 and is *compact*, then it is possible to find a *finite* number of open intervals I_k, $k = 1, \ldots, N$, that satisfy the two conditions (i) and (ii) above.

We can prove the characterization of Riemann integrable functions in terms of their discontinuities.

Theorem 1.7 *A bounded function f on $[a, b]$ is integrable if and only if its set of discontinuities has measure* 0.

[1] A systematic study of the measure of sets arises in the theory of Lebesgue integration, which is taken up in Book III.

We write $J = [a, b]$ and $I(c, r) = (c - r, c + r)$ for the open interval centered at c of radius $r > 0$. Define the **oscillation of** f **on** $I(c, r)$ by

$$\operatorname{osc}(f, c, r) = \sup |f(x) - f(y)|$$

where the supremum is taken over all $x, y \in J \cap I(c, r)$. This quantity exists since f is bounded. Define the **oscillation of** f **at** c by

$$\operatorname{osc}(f, c) = \lim_{r \to 0} \operatorname{osc}(f, c, r).$$

This limit exists because $\operatorname{osc}(f, c, r)$ is ≥ 0 and a decreasing function of r. The point is that f is continuous at c if and only if $\operatorname{osc}(f, c) = 0$. This is clear from the definitions. For each $\epsilon > 0$ we define a set A_ϵ by

$$A_\epsilon = \{c \in J : \operatorname{osc}(f, c) \geq \epsilon\}.$$

Having done that, we see that the set of points in J where f is discontinuous is simply $\bigcup_{\epsilon > 0} A_\epsilon$. This is an important step in the proof of our theorem.

Lemma 1.8 *If $\epsilon > 0$, then the set A_ϵ is closed and therefore compact.*

Proof. The argument is simple. Suppose $c_n \in A_\epsilon$ converges to c and assume that $c \notin A_\epsilon$. Write $\operatorname{osc}(f, c) = \epsilon - \delta$ where $\delta > 0$. Select r so that $\operatorname{osc}(f, c, r) < \epsilon - \delta/2$, and choose n with $|c_n - c| < r/2$. Then $\operatorname{osc}(f, c_n, r/2) < \epsilon$ which implies $\operatorname{osc}(f, c_n) < \epsilon$, a contradiction.

We are now ready to prove the first part of the theorem. Suppose that the set $\mathcal{D}$ of discontinuities of f has measure 0, and let $\epsilon > 0$. Since $A_\epsilon \subset \mathcal{D}$, we can cover A_ϵ by a finite number of open intervals, say $I_1, \ldots, I_N$, whose total length is $< \epsilon$. The complement of this union I of intervals is compact, and around each point z in this complement we can find an interval F_z with $\sup_{x,y \in F_z} |f(x) - f(y)| \leq \epsilon$, simply because $z \notin A_\epsilon$. We may now choose a finite subcovering of $\cup_{z \in I^c} I_z$, which we denote by $I_{N+1}, \ldots, I_{N'}$. Now, taking all the end points of the intervals $I_1, I_2, \ldots, I_{N'}$ we obtain a partition P of $[a, b]$ with

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) \leq 2M \sum_{j=1}^{N} |I_j| + \epsilon(b - a) \leq C\epsilon.$$

Hence f is integrable on $[a, b]$, as was to be shown.

Conversely, suppose that f is integrable on $[a, b]$, and let $\mathcal{D}$ be its set of discontinuities. Since $\mathcal{D}$ equals $\cup_{n=1}^{\infty} A_{1/n}$, it suffices to prove

that each $A_{1/n}$ has measure 0. Let $\epsilon > 0$ and choose a partition $P = \{x_0, x_1, \ldots, x_N\}$ so that $\mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon/n$. Then, if $A_{1/n}$ intersects $I_j = (x_{j-1}, x_j)$ we must have $\sup_{x \in I_j} f(x) - \inf_{x \in I_j} f(x) \geq 1/n$, and this shows that

$$\frac{1}{n} \sum_{\{j : I_j \cap A_{1/n} \neq \emptyset\}} |I_j| \leq \mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon/n.$$

So by taking intervals intersecting $A_{1/n}$ and making them slightly larger, we can cover $A_{1/n}$ with open intervals of total length $\leq 2\epsilon$. Therefore $A_{1/n}$ has measure 0, and we are done.

Note that incidentally, this gives another proof that fg is integrable whenever f and g are.

2 Multiple integrals

We assume that the reader is familiar with the standard theory of multiple integrals of functions defined on bounded sets. Here, we give a quick review of the main definitions and results of this theory. Then, we describe the notion of "improper" multiple integration where the range of integration is extended to all of $\mathbb{R}^d$. This is relevant to our study of the Fourier transform. In the spirit of Chapters 5 and 6, we shall define the integral of functions that are continuous and satisfy an adequate decay condition at infinity.

Recall that the vector space $\mathbb{R}^d$ consists of all d-tuples of real numbers $x = (x_1, x_2, \ldots, x_d)$ with $x_j \in \mathbb{R}$, where addition and multiplication by scalars are defined componentwise.

2.1 The Riemann integral in $\mathbb{R}^d$

Definitions

The notion of Riemann integration on a rectangle $R \subset \mathbb{R}^d$ is an immediate generalization of the notion of Riemann integration on an interval $[a, b] \subset \mathbb{R}$. We restrict our attention to continuous functions; these are always integrable.

By a **closed rectangle** in $\mathbb{R}^d$, we mean a set of the form

$$R = \{a_j \leq x_j \leq b_j \ : 1 \leq j \leq d\}$$

where $a_j, b_j \in \mathbb{R}$ for $1 \leq j \leq n$. In other words, R is the product of the one-dimensional intervals $[a_j, b_j]$:

$$R = [a_1, b_1] \times \cdots \times [a_d, b_d].$$

If P_j is a partition of the closed interval $[a_j, b_j]$, then we call $P = (P_1, \ldots, P_d)$ a **partition** of R; and if S_j is a subinterval of the partition P_j, then $S = S_1 \times \cdots \times S_d$ is a **subrectangle** of the partition P. The volume $|S|$ of a subrectangle S is naturally given by the product of the length of its sides $|S| = |S_1| \times \cdots \times |S_d|$, where $|S_j|$ denotes the length of the interval S_j.

We are now ready to define the notion of integral over R. Given a bounded real-valued function f defined on R and a partition P, we define the upper and lower sums of f with respect to P by

$$\mathcal{U}(P, f) = \sum [\sup_{x \in S} f(x)] \, |S| \quad \text{and} \quad \mathcal{L}(P, f) = \sum [\inf_{x \in S} f(x)] \, |S|,$$

where the sums are taken over all subrectangles of the partition P. These definitions are direct generalizations of the analogous notions in one dimension.

A partition $P' = (P'_1, \ldots, P'_d)$ is a **refinement** of $P = (P_1, \ldots, P_d)$ if each P'_j is a refinement of P_j. Arguing with these refinements as we did in the one-dimensional case, we see that if we define

$$U = \inf_P \ \mathcal{U}(P, f) \quad \text{and} \quad L = \sup_P \ \mathcal{L}(P, f),$$

then both U and L exist, are finite, and $U \geq L$. We say that f is **Riemann integrable** on R if for every $\epsilon > 0$ there exists a partition P so that

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon.$$

This implies that $U = L$, and this common value, which we shall denote by either

$$\int_R f(x_1, \ldots, x_d) \, dx_1 \cdots dx_d, \qquad \int_R f(x) \, dx, \qquad \text{or} \qquad \int_R f,$$

is by definition the integral of f over R. If f is complex-valued, say $f(x) = u(x) + iv(x)$, where u and v are real-valued, we naturally define

$$\int_R f(x) \, dx = \int_R u(x) \, dx + i \int_R v(x) \, dx.$$

In the results that follow, we are primarily interested in continuous functions. Clearly, if f is continuous on a closed rectangle R then f is integrable since it is uniformly continuous on R. Also, we note that if f is continuous on, say, a closed ball B, then we may define its integral

over B in the following way: if g is the extension of f defined by $g(x) = 0$ if $x \notin B$, then g is integrable on any rectangle R that contains B, and we may set

$$\int_B f(x)\,dx = \int_R g(x)\,dx.$$

2.2 Repeated integrals

The fundamental theorem of calculus allows us to compute many one dimensional integrals, since it is possible in many instances to find an antiderivative for the integrand. In $\mathbb{R}^d$, this permits the calculation of multiple integrals, since a d-dimensional integral actually reduces to d one-dimensional integrals. A precise statement describing this fact is given by the following.

Theorem 2.1 *Let f be a continuous function defined on a closed rectangle $R \subset \mathbb{R}^d$. Suppose $R = R_1 \times R_2$ where $R_1 \subset \mathbb{R}^{d_1}$ and $R_2 \subset \mathbb{R}^{d_2}$ with $d = d_1 + d_2$. If we write $x = (x_1, x_2)$ with $x_i \in \mathbb{R}^{d_i}$, then $F(x_1) = \int_{R_2} f(x_1, x_2)\,dx_2$ is continuous on R_1, and we have*

$$\int_R f(x)\,dx = \int_{R_1} \left(\int_{R_2} f(x_1, x_2)\,dx_2 \right) dx_1.$$

Proof. The continuity of F follows from the uniform continuity of f on R and the fact that

$$|F(x_1) - F(x_1')| \leq \int_{R_2} |f(x_1, x_2) - f(x_1', x_2)|\,dx_2.$$

To prove the identity, let P_1 and P_2 be partitions of R_1 and R_2, respectively. If S and T are subrectangles in P_1 and P_2, respectively, then the key observation is that

$$\sup_{S\times T} f(x_1, x_2) \geq \sup_{x_1 \in S} \left(\sup_{x_2 \in T} f(x_1, x_2) \right)$$

and

$$\inf_{S\times T} f(x_1, x_2) \leq \inf_{x_1 \in S} \left(\inf_{x_2 \in T} f(x_1, x_2) \right).$$

Then,

$$\begin{aligned}\mathcal{U}(P,f) &= \sum_{S,T} [\sup_{S\times T} f(x_1,x_2)]\,|S\times T| \\ &\geq \sum_S \sum_T \sup_{x_1\in S} [\sup_{x_2\in T} f(x_1,x_2)]\,|T|\times|S| \\ &\geq \sum_S \sup_{x_1\in S}\left(\int_{R_2} f(x_1,x_2)\,dx_2\right)|S| \\ &\geq \mathcal{U}\left(P_1, \int_{R_2} f(x_1,x_2)\,dx_2\right).\end{aligned}$$

Arguing similarly for the lower sums, we find that

$$\mathcal{L}(P,f) \leq \mathcal{L}(P_1, \int_{R_2} f(x_1,x_2)\,dx_2) \leq \mathcal{U}(P_1, \int_{R_2} f(x_1,x_2)\,dx_2) \leq \mathcal{U}(P,f),$$

and the theorem follows from these inequalities.

Repeating this argument, we find as a corollary that if f is continuous on the rectangle $R \subset \mathbb{R}^d$ given by $R = [a_1, b_1] \times \cdots [a_d, b_d]$, then

$$\int_R f(x)\,dx = \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} \cdots \left(\int_{a_d}^{b_d} f(x_1,\ldots,x_d)\,dx_d\right)\ldots dx_2\right)dx_1,$$

where the right-hand side denotes d-iterates of one-dimensional integrals. It is also clear from the theorem that we can interchange the order of integration in the repeated integral as desired.

2.3 The change of variables formula

A diffeomorphism of class C^1, $g : A \to B$, is a mapping that is continuously differentiable, invertible, and whose inverse $g^{-1} : B \to A$ is also continuously differentiable. We denote by Dg the Jacobian or derivative of g. Then, the change of variables formula says the following.

Theorem 2.2 *Suppose A and B are compact subsets of $\mathbb{R}^d$ and $g : A \to B$ is a diffeomorphism of class C^1. If f is continuous on B, then*

$$\int_{g(A)} f(x)\,dx = \int_A f(g(y))\,|\det(Dg)(y)|\,dy.$$

The proof of this theorem consists first of an analysis of the special situation when g is a linear transformation L. In this case, if R is a rectangle, then

$$|g(R)| = |\det(L)|\,|R|,$$

which explains the term $|\det(Dg)|$. Indeed, this term corresponds to the new infinitesimal element of volume after the change of variables.

2.4 Spherical coordinates

An important application of the change of variables formula is to the case of polar coordinates in $\mathbb{R}^2$, spherical coordinates in $\mathbb{R}^3$, and their generalization in $\mathbb{R}^d$. These are particularly important when the function, or set we are integrating over, exhibit some rotational (or spherical) symmetries. The cases $d = 2$ and $d = 3$ were given in Chapter 6. More generally, the spherical coordinates system in $\mathbb{R}^d$ is given by $x = g(r, \theta_1, \ldots, \theta_{d-1})$ where

$$\begin{cases} x_1 &= r \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{d-2}\cos\theta_{d-1}, \\ x_2 &= r \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{d-2}\sin\theta_{d-1}, \\ \vdots & \\ x_{d-1} &= r\sin\theta_1\sin\theta_2, \\ x_d &= r\cos\theta_1, \end{cases}$$

with $0 \leq \theta_i \leq \pi$ for $1 \leq i \leq d-2$ and $0 \leq \theta_{d-1} \leq 2\pi$. The determinant of the Jacobian of this transformation is given by

$$r^{d-1} \sin^{d-2}\theta_1 \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2}.$$

Any point in $x \in \mathbb{R}^d - \{0\}$ can be written uniquely as $r\gamma$ with $\gamma \in S^{d-1}$ the unit sphere in $\mathbb{R}^d$. If we define

$$\int_{S^{d-1}} f(\gamma)\, d\sigma(\gamma) =$$

$$\int_0^{\pi}\int_0^{\pi}\cdots\int_0^{2\pi} f(g(r,\theta)) \sin^{d-2}\theta_1 \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2}\, d\theta_{d-1}\cdots d\theta_1,$$

then we see that if $B(0, N)$ denotes the ball of radius N centered at the origin, then

$$\int_{B(0,N)} f(x)\, dx = \int_{S^{d-1}}\int_0^N f(r\gamma)\, r^{d-1}\, dr\, d\sigma(\gamma). \tag{3}$$

In fact, we define the **area** of the unit sphere $S^{d-1} \subset \mathbb{R}^d$ as

$$\omega_d = \int_{S^{d-1}} d\sigma(\gamma).$$

An important application of spherical coordinates is to the calculation of the integral $\int_{A(R_1,R_2)} |x|^{\lambda}\, dx$, where $A(R_1, R_2)$ denotes the annulus

$A(R_1, R_2) = \{R_1 \leq |x| \leq R_2\}$ and $\lambda \in \mathbb{R}$. Applying polar coordinates, we find

$$\int_{A(R_1,R_2)} |x|^\lambda \, dx = \int_{S^{d-1}} \int_{R_1}^{R_2} r^{\lambda+d-1} \, dr d\sigma(\gamma).$$

Therefore

$$\int_{A(R_1,R_2)} |x|^\lambda \, dx = \begin{cases} \frac{\omega_d}{\lambda+d}[R_2^{\lambda+d} - R_1^{\lambda+d}] & \text{if } \lambda \neq -d, \\ \omega_d[\log(R_2) - \log(R_1)] & \text{if } \lambda = -d. \end{cases}$$

3 Improper integrals. Integration over $\mathbb{R}^d$

Most of the theorems we just discussed extend to functions integrated over all of $\mathbb{R}^d$ once we impose some decay at infinity on the functions we integrate.

3.1 Integration of functions of moderate decrease

For each fixed $N > 0$ consider the closed cube in $\mathbb{R}^d$ centered at the origin with sides parallel to the axis, and of side length N: $Q_N = \{|x_j| \leq N/2 : 1 \leq j \leq d\}$. Let f be a continuous function on $\mathbb{R}^d$. If the limit

$$\lim_{N \to \infty} \int_{Q_N} f(x) \, dx$$

exists, we denote it by

$$\int_{\mathbb{R}^d} f(x) \, dx.$$

We deal with a special class of functions whose integrals over $\mathbb{R}^d$ exist. A continuous function f on $\mathbb{R}^d$ is said to be of **moderate decrease** if there exists $A > 0$ such that

$$|f(x)| \leq \frac{A}{1 + |x|^{d+1}}.$$

Note that if $d = 1$ we recover the definition given in Chapter 5. An important example of a function of moderate decrease in $\mathbb{R}$ is the Poisson kernel given by $\mathcal{P}_y(x) = \frac{1}{\pi} \frac{y}{x^2+y^2}$.

We claim that if f is of moderate decrease, then the above limit exists. Let $I_N = \int_{Q_N} f(x) \, dx$. Each I_N exists because f is continuous hence integrable. For $M > N$, we have

$$|I_M - I_N| \leq \int_{Q_M - Q_N} |f(x)| \, dx.$$

Now observe that the set $Q_M - Q_N$ is contained in the annulus $A(aN, bM) = \{aN \leq |x| \leq bM\}$, where a and b are constants that depend only on the dimension d. This is because the cube Q_N is contained in the annulus $N/2 \leq |x| \leq N\sqrt{d}/2$, so that we can take $a = 1/2$ and $b = \sqrt{d}/2$. Therefore, using the fact that f is of moderate decrease yields

$$|I_M - I_N| \leq A \int_{aN \leq |x| \leq bM} |x|^{-d-1}\, dx.$$

Now putting $\lambda = -d-1$ in the calculation of the integral of the previous section, we find that

$$|I_M - I_N| \leq C \left(\frac{1}{aN} - \frac{1}{bM} \right).$$

So if f is of moderate decrease, we conclude that $\{I_N\}_{N=1}^{\infty}$ is a Cauchy sequence, and therefore $\int_{\mathbb{R}^d} f(x)\, dx$ exists.

Instead of the rectangles Q_N, we could have chosen the balls B_N centered at the origin and of radius N. Then, if f is of moderate decrease, the reader should have no difficulties proving that $\lim_{N\to\infty} \int_{B_N} f(x)\, dx$ exists, and that this limit equals $\lim_{N\to\infty} \int_{Q_N} f(x)\, dx$.

Some elementary properties of the integrals of functions of moderate decrease are summarized in Chapter 6.

3.2 Repeated integrals

In Chapters 5 and 6 we claimed that the multiplication formula held for functions of moderate decrease. This required an appropriate interchange of integration. Similarly for operators defined in terms of convolutions (with the Poisson kernel for example).

We now justify the necessary formula for iterated integrals. We only consider the case $d = 2$, although the reader will have no difficulty extending this result to arbitrary dimensions.

Theorem 3.1 *Suppose f is continuous on $\mathbb{R}^2$ and of moderate decrease. Then*

$$F(x_1) = \int_{\mathbb{R}} f(x_1, x_2)\, dx_2$$

is of moderate decrease on $\mathbb{R}$, and

$$\int_{\mathbb{R}^2} f(x)\, dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, x_2)\, dx_2 \right) dx_1.$$

Proof. To see why F is of moderate decrease, note first that

$$|F(x_1)| \leq \int_{\mathbb{R}} \frac{A\,dx_2}{1+(x_1^2+x_2^2)^{3/2}} \leq \int_{|x_2|\leq|x_1|} + \int_{|x_2|\geq|x_1|}.$$

In the first integral, we observe that the integrand is $\leq A/(1+|x_1|^3)$, so

$$\int_{|x_2|\leq|x_1|} \frac{A\,dx_2}{1+(x_1^2+x_2^2)^{3/2}} \leq \frac{A}{1+|x_1|^3}\int_{|x_2|\leq|x_1|} dx_2 \leq \frac{A'}{1+|x_1|^2}.$$

For the second integral, we have

$$\int_{|x_2|\geq|x_1|} \frac{A\,dx_2}{1+(x_1^2+x_2^2)^{3/2}} \leq A''\int_{|x_2|\geq|x_1|} \frac{dx_2}{1+|x_2|^3} \leq \frac{A'''}{|x_1|^2},$$

thus F is of moderate decrease. In fact, this argument together with Theorem 2.1 shows that F is the uniform limit of continuous functions, thus is also continuous.

To establish the identity we simply use an approximation and Theorem 2.1 over finite rectangles. Write S^c to denote the complement of a set S. Given $\epsilon > 0$ choose N so large that

$$\left| \int_{\mathbb{R}^2} f(x_1,x_2)\,dx_1dx_2 - \int_{I_N\times I_N} f(x_1,x_2)\,dx_1dx_2 \right| < \epsilon,$$

where $I_N = [-N,N]$. Now we know that

$$\int_{I_N\times I_N} f(x_1,x_2)\,dx_1dx_2 = \int_{I_N}\left(\int_{I_N} f(x_1,x_2)\,dx_2\right)dx_1.$$

But this last iterated integral can be written as

$$= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(x_1,x_2)\,dx_2\right)dx_1 - \int_{I_N^c}\left(\int_{\mathbb{R}} f(x_1,x_2)\,dx_2\right)dx_1 - \int_{I_N}\left(\int_{I_N^c} f(x_1,x_2)\,dx_2\right)dx_1.$$

We can now estimate

$$\begin{aligned}\left|\int_{I_N}\left(\int_{I_N^c} f(x_1,x_2)\,dx_2\right)dx_1\right| &\leq O\left(\frac{1}{N^2}\right) \\ &\quad + C\int_{1\leq|x_1|\leq N}\left(\int_{|x_2|\geq N} \frac{dx_2}{(|x_1|+|y_1|)^3}\right)dx_1 \\ &\leq O\left(\frac{1}{N}\right).\end{aligned}$$

A similar argument shows that

$$\left| \int_{I_N^c} \left(\int_{\mathbb{R}} f(x_1, x_2)\, dx_2 \right) dx_1 \right| \leq \frac{C}{N}.$$

Therefore, we can find N so large that

$$\left| \int_{I_N \times I_N} f(x_1, x_2)\, dx_1 dx_2 - \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, x_2)\, dx_2 \right) dx_1 \right| < \epsilon,$$

and we are done.

3.3 Spherical coordinates

In $\mathbb{R}^d$, spherical coordinates are given by $x = r\gamma$, where $r \geq 0$ and γ belongs to the unit sphere S^{d-1}. If f is of moderate decrease, then for each fixed $\gamma \in S^{d-1}$, the function of f given by $f(r\gamma)r^{d-1}$ is also of moderate decrease on $\mathbb{R}$. Indeed, we have

$$\left| f(r\gamma) r^{d-1} \right| \leq A \frac{r^{d-1}}{1 + |r\gamma|^{d+1}} \leq \frac{B}{1 + r^2}.$$

As a result, by letting $R \to \infty$ in (3) we obtain the formula

$$\int_{\mathbb{R}^d} f(x)\, dx = \int_{S^{d-1}} \int_0^\infty f(r\gamma)\, r^{d-1}\, dr\, d\sigma(\gamma).$$

As a consequence, if we combine the fact that

$$\int_{\mathbb{R}^d} f(R(x))\, dx = \int_{\mathbb{R}^d} f(x)\, dx,$$

whenever R is a rotation, with the identity (3), then we obtain that

$$\int_{S^{d-1}} f(R(\gamma))\, d\sigma(\gamma) = \int_{S^{d-1}} f(\gamma)\, d\sigma(\gamma)\,. \tag{4}$$

Notes and References

Seeley [29] gives an elegant and brief introduction to Fourier series and the Fourier transform. The authoritative text on Fourier series is Zygmund [36]. For further applications of Fourier analysis to a variety of other topics, see Dym and McKean [8] and Körner [21]. The reader should also consult the book by Kahane and Lemarié-Rieusset [20], which contains many historical facts and other results related to Fourier series.

Chapter 1

The citation is taken from a letter of Fourier to an unknown correspondent (probably Lagrange), see Herivel [15].

More facts about the early history of Fourier series can be found in Sections I-III of Riemann's memoir [27].

Chapter 2

The quote is a translation of an excerpt in Riemann's paper [27].

For a proof of Littlewood's theorem (Problem 3), as well as other related "Tauberian theorems," see Chapter 7 in Titchmarsh [32].

Chapter 3

The citation is a translation of a passage in Dirichlet's memoir [6].

Chapter 4

The quote is translated from Hurwitz [17].

The problem of a ray of light reflecting inside a square is discussed in Chapter 23 of Hardy and Wright [13].

The relationship between the diameter of a curve and Fourier coefficients (Problem 1) is explored in Pfluger [26].

Many topics concerning equidistribution of sequences, including the results in Problems 2 and 3, are taken up in Kuipers and Niederreiter [22].

Chapter 5

The citation is a free translation of a passage in Schwartz [28].

For topics in finance, see Duffie [7], and in particular Chapter 5 for the Black-Scholes theory (Problems 1 and 2).

The results in Problems 4, 5, and 6 are worked out in John [19] and Widder [34].

For Problem 7, see Chapter 2 in Wiener [35].

The original proof of the nowhere differentiability of f_1 (Problem 8) is in Hardy [12].

Chapter 6

The quote is an excerpt from Cormack's Nobel Prize lecture [5].

More about the wave equation, as well as the results in Problems 3, 4, and 5 can be found in Chapter 5 of Folland [9].

A discussion of the relationship between rotational symmetry, the Fourier transform, and Bessel functions is in Chapter 4 of Stein and Weiss [31].

For more on the Radon transform, see Chapter 1 in John [18], Helgason [14], and Ludwig [25].

Chapter 7

The citation is taken from Bingham and Tukey [2].

Proofs of the structure theorem for finite abelian groups (Problem 2) can be found in Chapter 2 of Herstein [16], Chapter 2 in Lang [23], or Chapter 104 in Körner [21].

For Problem 4, see Andrews [1], which contains a short proof.

Chapter 8

The citation is from Bochner [3].

For more on the divisor function, see Chapter 18 in Hardy and Wright [13].

Another "elementary" proof that $L(1, \chi) \neq 0$ can be found in Chapter 3 of Gelfond and Linnik [11].

An alternate proof that $L(1, \chi) \neq 0$ based on algebraic number theory is in Weyl [33]. Also, two other analytic variants of the proof that $L(1, \chi) \neq 0$ can be found in Chapter 109 in Körner [21] and Chapter 6 in Serre [30]. See also the latter reference for Problems 3 and 4.

Appendix

Further details about the results on integration reviewed in the appendix can be found in Folland [10] (Chapter 4), Buck [4] (Chapter 4), or Lang [24] (Chapter 20).

Bibliography

[1] G. E. Andrews. *Number theory.* Dover Publications, New York, 1994. Corrected reprint of the 1971 originally published by W.B. Saunders Company.

[2] C. Bingham and J.W. Tukey. Fourier methods in the frequency analysis of data. *The Collected Works of John W Tukey*, Volume II Time Series: 1965-1984(Wadsworth Advanced Books & Software), 1984.

[3] S. Bochner. *The role of Mathematics in the Rise of Science.* Princeton University Press, Princeton, NJ, 1966.

[4] R. C. Buck. *Advanced Calculus.* McGraw-Hill, New York, third edition, 1978.

[5] A. M. Cormack. *Nobel Prize in Physiology and Medicine Lecture*, volume Volume 209. Science, 1980.

[6] G. L. Dirichlet. Sur la convergence des séries trigonometriques qui servent à representer une fonction arbitraire entre des limites données. *Crelle, Journal für die reine angewandte Mathematik*, 4:157–169, 1829.

[7] D. Duffie. *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton, NJ, 2001.

[8] H. Dym and H. P. McKean. *Fourier Series and Integrals.* Academic Press, New York, 1972.

[9] G. B. Folland. *Introduction to Partial Differential Equations.* Princeton University Press, Princeton, NJ, 1995.

[10] G. B. Folland. *Advanced Calculus.* Prentice Hall, Englewood Cliffs, NJ, 2002.

[11] A. O. Gelfond and Yu. V. Linnik. *Elementary Methods in Analytic Number Theory.* Rand McNally & Compagny, Chicago, 1965.

[12] G. H. Hardy. Weierstrass's non-differentiable function. *Transactions, American Mathematical Society*, 17:301–325, 1916.

[13] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, London, fifth edition, 1979.

[14] S. Helgason. The Radon transform on Euclidean spaces, compact two-point homogeneous spaces and Grassman manifolds. *Acta. Math.*, 113:153–180, 1965.

[15] J. Herivel. *Joseph Fourier The Man and the Physicist.* Clarendon Press, Oxford, 1975.

[16] I. N. Herstein. *Abstract Algebra.* Macmillan, New York, second edition, 1990.

[17] A. Hurwitz. Sur quelques applications géometriques des séries de Fourier. *Annales de l'Ecole Normale Supérieure*, 19(3):357–408, 1902.

[18] F. John. *Plane Waves and Spherical Mean Applied to Partial Differential Equations.* Interscience Publishers, New York, 1955.

[19] F. John. *Partial Differential Equations.* Springer-Verlag, New York, fourth edition, 1982.

[20] J.P. Kahane and P. G. Lemarié-Rieusset. *Séries de Fourier et ondelettes.* Cassini, Paris, 1998. English version: Gordon & Breach, 1995.

[21] T. W. Körner. *Fourier Analysis.* Cambridge University Press, Cambridge, UK, 1988.

[22] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences.* Wiley, New York, 1974.

[23] S. Lang. *Undergraduate Algebra.* Springer-Verlag, New York, second edition, 1990.

[24] S. Lang. *Undergraduate Analysis.* Springer-Verlag, New York, second edition, 1997.

[25] D. Ludwig. The Radon transform on Euclidean space. *Comm. Pure Appl. Math.*, 19:49–81, 1966.

[26] A. Pfluger. On the diameter of planar curves and Fourier coefficients. *Colloquia Mathematica Societatis János Bolyai, Functions, series, operators*, 35:957–965, 1983.

[27] B. Riemann. Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe. *Habilitation an der Universität zu Göttingen*, 1854. Collected Works, Springer Verlag, New York, 1990.

[28] L. Schwartz. *Théorie des distributions*, volume Volume I. Hermann, Paris, 1950.

[29] R. T. Seeley. *An Introduction to Fourier Series and Integrals.* W. A. Benjamin, New York, 1966.

[30] J.P. Serre. *A course in Arithmetic.* GTM 7. Springer Verlag, New York, 1973.

[31] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, Princeton, NJ, 1971.

[32] E. C. Titchmarsh. *The Theory of Functions.* Oxford University Press, London, second edition, 1939.

[33] H. Weyl. *Algebraic Theory of Numbers*, volume Volume 1 of *Annals of Mathematics Studies.* Princeton University Press, Princeton, NJ, 1940.

[34] D. V. Widder. *The Heat Equation.* Academic Press, New York, 1975.

[35] N. Wiener. *The Fourier Integral and Certain of its Applications.* Cambridge University Press, Cambridge, UK, 1933.

[36] A. Zygmund. *Trigonometric Series*, volume Volumes I and II. Cambridge University Press, Cambridge, UK, second edition, 1959. Reprinted 1993.

Symbol Glossary

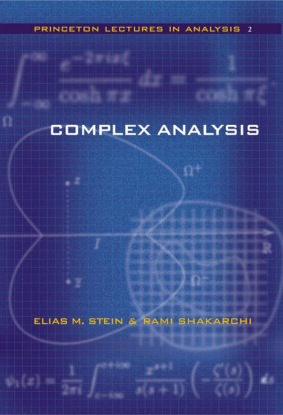
The page numbers on the right indicate the first time the symbol or notation is defined or used. As usual, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ denote the integers, the rationals, the reals, and the complex numbers respectively.

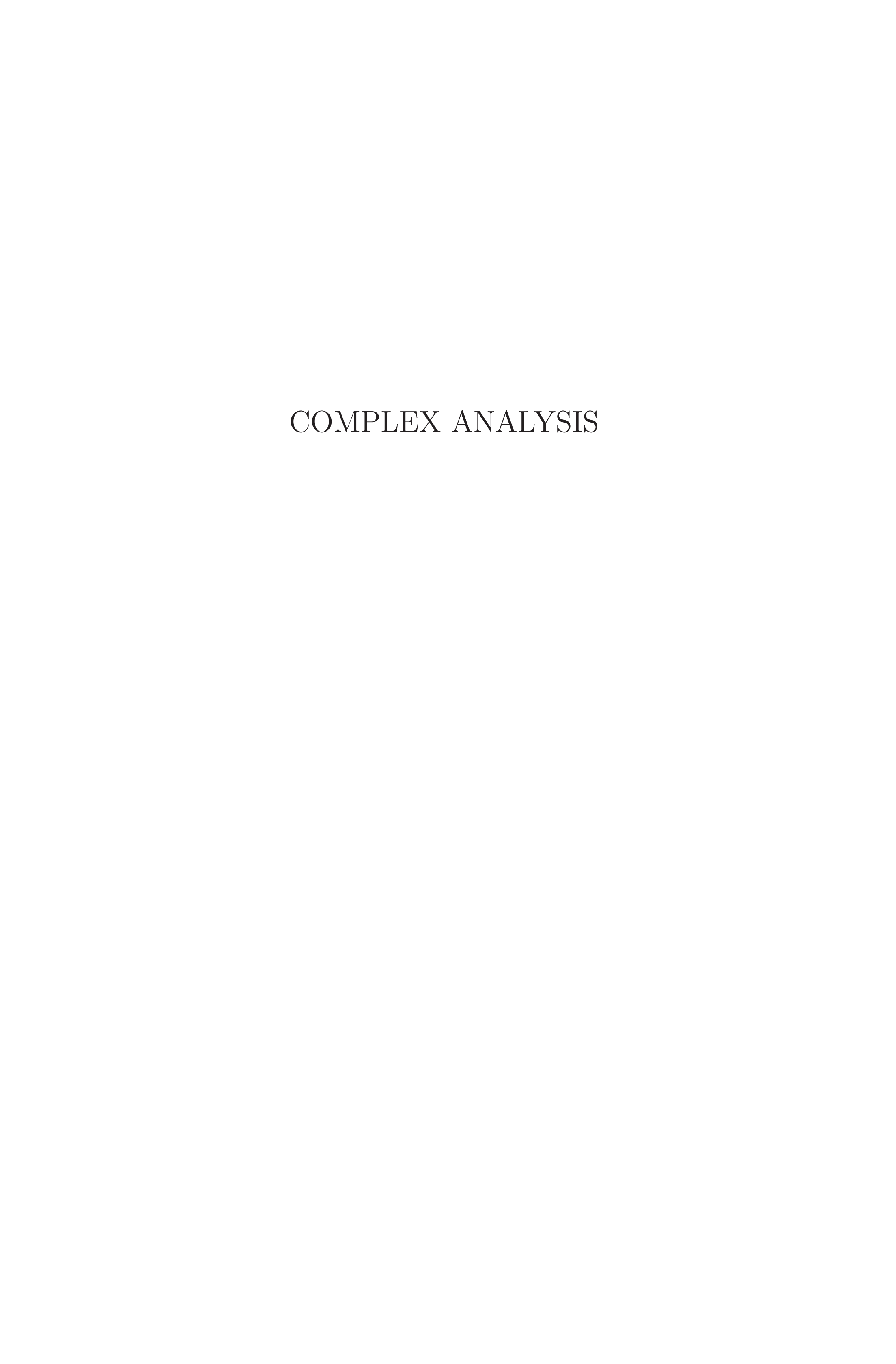
$\triangle$	Laplacian	20, 185
$\lvert z\rvert$, $\overline{z}$	Absolute value and complex conjugate	22
e^z	Complex exponential	24
$\sinh x$, $\cosh x$	Hyperbolic sine and hyperbolic cosine	28
$\hat{f}(n)$, a_n	Fourier coefficient	34
$f(\theta) \sim \sum a_n e^{in\theta}$	Fourier series	34
$S_N(f)$	Partial sum of a Fourier series	35
D_N, $\mathcal{D}_R$, $\tilde{D}_N$, D_N^*	Dirichlet kernel, conjugate, and modified	37, 95, 165
P_r, $\mathcal{P}_y$, $\mathcal{P}_y^{(d)}$	Poisson kernels	37, 149, 210
O, o	Big O and little o notation	42, 62
C^k	Functions that are k times differentiable	44
$f * g$	Convolution	44, 139, 184, 239
σ_N, $\sigma_N(f)$	Cesàro mean	52, 53
F_N, $\mathcal{F}_R$	Fejér kernel	53, 163
$A(r)$, $A_r(f)$	Abel mean	54, 55
$\chi_{[a,b]}$	Characteristic function	61
$f(\theta^+)$, $f(\theta^-)$	One-sided limits at jump discontinuities	63
$\mathbb{R}^d$, $\mathbb{C}^d$	Euclidean spaces	71
$X \perp Y$	Orthogonal vectors	72
$\ell^2(\mathbb{Z})$	Square summable sequences	73
$\mathcal{R}$	Riemann integrable functions	75
$\zeta(s)$	Zeta function	98
$[x]$, $\langle x\rangle$	Integer and fractional parts	106
$\triangle_N$, $\sigma_{N,K}$, $\tilde{\triangle}_N$	Delayed means	114, 127, 174
H_t, $\mathcal{H}_t$, $\mathcal{H}_t^{(d)}$	Heat kernels	120, 146, 209
$\mathcal{M}(\mathbb{R})$	Space of functions of moderate decrease on $\mathbb{R}$	131

Symbol	Meaning	Pages
$\hat{f}(\xi)$	Fourier transform	134, 181
$\mathcal{S}, \mathcal{S}(\mathbb{R}), \mathcal{S}(\mathbb{R}^d)$	Schwartz space	134, 180
$\mathbb{R}^2_+$, $\overline{\mathbb{R}^2_+}$	Upper half-plane and its closure	149
$\vartheta(s)$, $\Theta(z\|\tau)$	Theta functions	155, 156
$\Gamma(s)$	Gamma function	165
$\\|x\\|$, $\|x\|$; (x, y), $x \cdot y$	Norm and inner product in $\mathbb{R}^d$	71, 176
x^α, $\|\alpha\|$, $\left(\frac{\partial}{\partial x}\right)^\alpha$	Monomial, its order, and differential operator	176
S^1, S^2, S^{d-1}	Unit circle in $\mathbb{R}^2$, and unit spheres in $\mathbb{R}^3$, $\mathbb{R}^d$	179, 180
M_t, $\widetilde{M}_t$	Spherical mean	194, 216
J_n	Bessel function	197, 213
$\mathcal{P}$, $\mathcal{P}_{t,\gamma}$	Plane	202
$\mathcal{R}$, $\mathcal{R}^*$	Radon and dual Radon transforms	203, 205
A_d, V_d	Area and volume of the unit sphere in $\mathbb{R}^d$	208
$\mathbb{Z}(N)$	Group of Nth roots of unity	219
$\mathbb{Z}/N\mathbb{Z}$	Group of integers modulo N	221
G, $\|G\|$	Abelian group and its order	226, 228
$G \approx H$	Isomorphic groups	227
$G_1 \times G_2$	Direct product of groups	228
$\mathbb{Z}^*(q)$	Group of units modulo q	227, 229, 244
$\hat{G}$	Dual group of G	231
$a\|b$	a divides b	242
$\gcd(a, b)$	Greatest common divisor of a and b	242
$\varphi(q)$	Number of integers relatively prime to q	254
χ, χ_0	Dirichlet character, and trivial Dirichlet character	254
$L(s, \chi)$	Dirichlet L-function	256
$\log_1\left(\frac{1}{1-z}\right)$, $\log_2 L(s, \chi)$	Logarithms	258, 264
$d(k)$	Number of positive divisors of k	269

Index

Relevant items that also arise in Book I are listed in this index, preceeded by the numeral I.


PRINCETON LECTURES IN ANALYSIS 2
COMPLEX ANALYSIS
ELIAS M. STEIN & RAMI SHAKARCHI



COMPLEX ANALYSIS

Princeton Lectures in Analysis

I Fourier Analysis: An Introduction

II Complex Analysis

III Real Analysis:
Measure Theory, Integration, and
Hilbert Spaces

Princeton Lectures in Analysis

II

COMPLEX ANALYSIS

Elias M. Stein

&

Rami Shakarchi

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

Library of Congress Control Number 2005274996
ISBN 978-0-691-11385-2

British Library Cataloging-in-Publication Data is available

The publisher would like to acknowledge the authors of this volume for
providing the camera-ready copy from which this book was printed

Printed on acid-free paper. ∞

press.princeton.edu

Printed in the United States of America

5 7 9 10 8 6

To my grandchildren
Carolyn, Alison, Jason

E.M.S.

To my parents
Mohamed & Mireille
and my brother
Karim

R.S.

Foreword

Beginning in the spring of 2000, a series of four one-semester courses were taught at Princeton University whose purpose was to present, in an integrated manner, the core areas of analysis. The objective was to make plain the organic unity that exists between the various parts of the subject, and to illustrate the wide applicability of ideas of analysis to other fields of mathematics and science. The present series of books is an elaboration of the lectures that were given.

While there are a number of excellent texts dealing with individual parts of what we cover, our exposition aims at a different goal: presenting the various sub-areas of analysis not as separate disciplines, but rather as highly interconnected. It is our view that seeing these relations and their resulting synergies will motivate the reader to attain a better understanding of the subject as a whole. With this outcome in mind, we have concentrated on the main ideas and theorems that have shaped the field (sometimes sacrificing a more systematic approach), and we have been sensitive to the historical order in which the logic of the subject developed.

We have organized our exposition into four volumes, each reflecting the material covered in a semester. Their contents may be broadly summarized as follows:

I. Fourier series and integrals.

II. Complex analysis.

III. Measure theory, Lebesgue integration, and Hilbert spaces.

IV. A selection of further topics, including functional analysis, distributions, and elements of probability theory.

However, this listing does not by itself give a complete picture of the many interconnections that are presented, nor of the applications to other branches that are highlighted. To give a few examples: the elements of (finite) Fourier series studied in Book I, which lead to Dirichlet characters, and from there to the infinitude of primes in an arithmetic progression; the X-ray and Radon transforms, which arise in a number of

problems in Book I, and reappear in Book III to play an important role in understanding Besicovitch-like sets in two and three dimensions; Fatou's theorem, which guarantees the existence of boundary values of bounded holomorphic functions in the disc, and whose proof relies on ideas developed in each of the first three books; and the theta function, which first occurs in Book I in the solution of the heat equation, and is then used in Book II to find the number of ways an integer can be represented as the sum of two or four squares, and in the analytic continuation of the zeta function.

A few further words about the books and the courses on which they were based. These courses where given at a rather intensive pace, with 48 lecture-hours a semester. The weekly problem sets played an indispensable part, and as a result exercises and problems have a similarly important role in our books. Each chapter has a series of "Exercises" that are tied directly to the text, and while some are easy, others may require more effort. However, the substantial number of hints that are given should enable the reader to attack most exercises. There are also more involved and challenging "Problems"; the ones that are most difficult, or go beyond the scope of the text, are marked with an asterisk.

Despite the substantial connections that exist between the different volumes, enough overlapping material has been provided so that each of the first three books requires only minimal prerequisites: acquaintance with elementary topics in analysis such as limits, series, differentiable functions, and Riemann integration, together with some exposure to linear algebra. This makes these books accessible to students interested in such diverse disciplines as mathematics, physics, engineering, and finance, at both the undergraduate and graduate level.

It is with great pleasure that we express our appreciation to all who have aided in this enterprise. We are particularly grateful to the students who participated in the four courses. Their continuing interest, enthusiasm, and dedication provided the encouragement that made this project possible. We also wish to thank Adrian Banner and Jose Luis Rodrigo for their special help in running the courses, and their efforts to see that the students got the most from each class. In addition, Adrian Banner also made valuable suggestions that are incorporated in the text.

We wish also to record a note of special thanks for the following individuals: Charles Fefferman, who taught the first week (successfully launching the whole project!); Paul Hagelstein, who in addition to reading part of the manuscript taught several weeks of one of the courses, and has since taken over the teaching of the second round of the series; and Daniel Levine, who gave valuable help in proof-reading. Last but not least, our thanks go to Gerree Pecht, for her consummate skill in typesetting and for the time and energy she spent in the preparation of all aspects of the lectures, such as transparencies, notes, and the manuscript.

We are also happy to acknowledge our indebtedness for the support we received from the 250th Anniversary Fund of Princeton University, and the National Science Foundation's VIGRE program.

Elias M. Stein
Rami Shakarchi

Princeton, New Jersey
August 2002

Contents

Introduction

> ... In effect, if one extends these functions by allowing complex values for the arguments, then there arises a harmony and regularity which without it would remain hidden.
>
> *B. Riemann,* 1851

When we begin the study of complex analysis we enter a marvelous world, full of wonderful insights. We are tempted to use the adjectives magical, or even miraculous when describing the first theorems we learn; and in pursuing the subject, we continue to be astonished by the elegance and sweep of the results.

The starting point of our study is the idea of extending a function initially given for real values of the argument to one that is defined when the argument is complex. Thus, here the central objects are functions from the complex plane to itself

$$f : \mathbb{C} \to \mathbb{C},$$

or more generally, complex-valued functions defined on open subsets of $\mathbb{C}$. At first, one might object that nothing new is gained from this extension, since any complex number z can be written as $z = x + iy$ where $x, y \in \mathbb{R}$ and z is identified with the point (x, y) in $\mathbb{R}^2$.

However, everything changes drastically if we make a natural, but misleadingly simple-looking assumption on f: that it is differentiable in the complex sense. This condition is called **holomorphicity**, and it shapes most of the theory discussed in this book.

A function $f : \mathbb{C} \to \mathbb{C}$ is holomorphic at the point $z \in \mathbb{C}$ if the limit

$$\lim_{h \to 0} \frac{f(z+h) - f(z)}{h} \qquad (h \in \mathbb{C})$$

exists. This is similar to the definition of differentiability in the case of a real argument, except that we allow h to take *complex* values. The reason this assumption is so far-reaching is that, in fact, it encompasses a multiplicity of conditions: so to speak, one for each angle that h can approach zero.

Although one might now be tempted to prove theorems about holomorphic functions in terms of real variables, the reader will soon discover that complex analysis is a new subject, one which supplies proofs to the theorems that are proper to its own nature. In fact, the proofs of the main properties of holomorphic functions which we discuss in the next chapters are generally very short and quite illuminating.

The study of complex analysis proceeds along two paths that often intersect. In following the first way, we seek to understand the universal characteristics of holomorphic functions, without special regard for specific examples. The second approach is the analysis of some particular functions that have proved to be of great interest in other areas of mathematics. Of course, we cannot go too far along either path without having traveled some way along the other. We shall start our study with some general characteristic properties of holomorphic functions, which are subsumed by three rather miraculous facts:

1. Contour integration: If f is holomorphic in Ω, then for appropriate closed paths in Ω

$$\int_\gamma f(z)dz = 0.$$

2. Regularity: If f is holomorphic, then f is indefinitely differentiable.

3. Analytic continuation: If f and g are holomorphic functions in Ω which are equal in an arbitrarily small disc in Ω, then $f = g$ everywhere in Ω.

These three phenomena and other general properties of holomorphic functions are treated in the beginning chapters of this book. Instead of trying to summarize the contents of the rest of this volume, we mention briefly several other highlights of the subject.

- The zeta function, which is expressed as an infinite series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

and is initially defined and holomorphic in the half-plane $\text{Re}(s) > 1$, where the convergence of the sum is guaranteed. This function and its variants (the L-series) are central in the theory of prime numbers, and have already appeared in Chapter 8 of Book I, where

we proved Dirichlet's theorem. Here we will prove that ζ extends to a meromorphic function with a pole at $s = 1$. We shall see that the behavior of $\zeta(s)$ for $\text{Re}(s) = 1$ (and in particular that ζ does not vanish on that line) leads to a proof of the prime number theorem.

- The theta function

$$\Theta(z|\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{2\pi i n z},$$

which in fact is a function of the two complex variables z and τ, holomorphic for all z, but only for τ in the half-plane $\text{Im}(\tau) > 0$. On the one hand, when we fix τ, and think of Θ as a function of z, it is closely related to the theory of elliptic (doubly-periodic) functions. On the other hand, when z is fixed, Θ displays features of a modular function in the upper half-plane. The function $\Theta(z|\tau)$ arose in Book I as a fundamental solution of the heat equation on the circle. It will be used again in the study of the zeta function, as well as in the proof of certain results in combinatorics and number theory given in Chapters 6 and 10.

Two additional noteworthy topics that we treat are: the Fourier transform with its elegant connection to complex analysis via contour integration, and the resulting applications of the Poisson summation formula; also conformal mappings, with the mappings of polygons whose inverses are realized by the Schwarz-Christoffel formula, and the particular case of the rectangle, which leads to elliptic integrals and elliptic functions.

1 Preliminaries to Complex Analysis

> The sweeping development of mathematics during the last two centuries is due in large part to the introduction of complex numbers; paradoxically, this is based on the seemingly absurd notion that there are numbers whose squares are negative.
>
> *E. Borel,* 1952

This chapter is devoted to the exposition of basic preliminary material which we use extensively throughout of this book.

We begin with a quick review of the algebraic and analytic properties of complex numbers followed by some topological notions of sets in the complex plane. (See also the exercises at the end of Chapter 1 in Book I.)

Then, we define precisely the key notion of holomorphicity, which is the complex analytic version of differentiability. This allows us to discuss the Cauchy-Riemann equations, and power series.

Finally, we define the notion of a curve and the integral of a function along it. In particular, we shall prove an important result, which we state loosely as follows: if a function f has a primitive, in the sense that there exists a function F that is holomorphic and whose derivative is precisely f, then for any closed curve γ

$$\int_\gamma f(z)\,dz = 0.$$

This is the first step towards Cauchy's theorem, which plays a central role in complex function theory.

1 Complex numbers and the complex plane

Many of the facts covered in this section were already used in Book I.

1.1 Basic properties

A complex number takes the form $z = x + iy$ where x and y are real, and i is an imaginary number that satisfies $i^2 = -1$. We call x and y the

real part and the **imaginary part** of z, respectively, and we write

$$x = \mathrm{Re}(z) \quad \text{and} \quad y = \mathrm{Im}(z).$$

The real numbers are precisely those complex numbers with zero imaginary parts. A complex number with zero real part is said to be **purely imaginary**.

Throughout our presentation, the set of all complex numbers is denoted by $\mathbb{C}$. The complex numbers can be visualized as the usual Euclidean plane by the following simple identification: the complex number $z = x + iy \in \mathbb{C}$ is identified with the point $(x, y) \in \mathbb{R}^2$. For example, 0 corresponds to the origin and i corresponds to $(0, 1)$. Naturally, the x and y axis of $\mathbb{R}^2$ are called the **real axis** and **imaginary axis**, because they correspond to the real and purely imaginary numbers, respectively. (See Figure 1.)

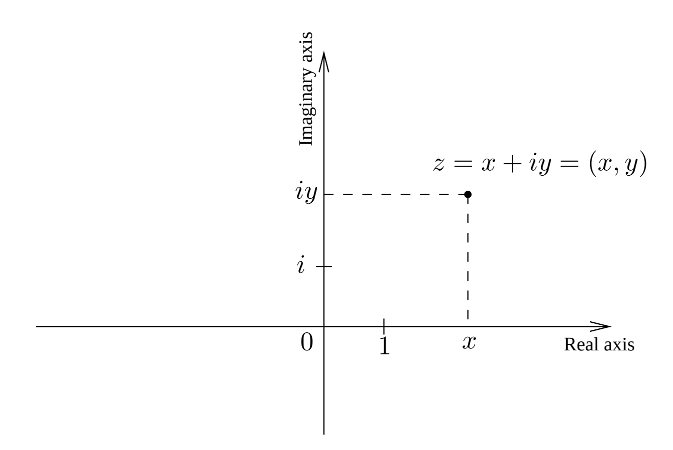


Figure 1. The complex plane

The natural rules for adding and multiplying complex numbers can be obtained simply by treating all numbers as if they were real, and keeping in mind that $i^2 = -1$. If $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, then

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2),$$

and also

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1 x_2 + i x_1 y_2 + i y_1 x_2 + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2). \end{aligned}$$

If we take the two expressions above as the definitions of addition and multiplication, it is a simple matter to verify the following desirable properties:

- Commutativity: $z_1 + z_2 = z_2 + z_1$ and $z_1 z_2 = z_2 z_1$ for all $z_1, z_2 \in \mathbb{C}$.
- Associativity: $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$; and $(z_1 z_2) z_3 = z_1(z_2 z_3)$ for $z_1, z_2, z_3 \in \mathbb{C}$.
- Distributivity: $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$ for all $z_1, z_2, z_3 \in \mathbb{C}$.

Of course, addition of complex numbers corresponds to addition of the corresponding vectors in the plane $\mathbb{R}^2$. Multiplication, however, consists of a rotation composed with a dilation, a fact that will become transparent once we have introduced the polar form of a complex number. At present we observe that multiplication by i corresponds to a rotation by an angle of $\pi/2$.

The notion of length, or absolute value of a complex number is identical to the notion of Euclidean length in $\mathbb{R}^2$. More precisely, we define the **absolute value** of a complex number $z = x + iy$ by

$$|z| = (x^2 + y^2)^{1/2},$$

so that $|z|$ is precisely the distance from the origin to the point (x, y). In particular, the triangle inequality holds:

$$|z + w| \leq |z| + |w| \qquad \text{for all } z, w \in \mathbb{C}.$$

We record here other useful inequalities. For all $z \in \mathbb{C}$ we have both $|\mathrm{Re}(z)| \leq |z|$ and $|\mathrm{Im}(z)| \leq |z|$, and for all $z, w \in \mathbb{C}$

$$||z| - |w|| \leq |z - w|.$$

This follows from the triangle inequality since

$$|z| \leq |z - w| + |w| \quad \text{and} \quad |w| \leq |z - w| + |z|.$$

The **complex conjugate** of $z = x + iy$ is defined by

$$\overline{z} = x - iy,$$

and it is obtained by a reflection across the real axis in the plane. In fact a complex number z is real if and only if $z = \overline{z}$, and it is purely imaginary if and only if $z = -\overline{z}$.

The reader should have no difficulty checking that

$$\mathrm{Re}(z) = \frac{z+\overline{z}}{2} \quad \text{and} \quad \mathrm{Im}(z) = \frac{z-\overline{z}}{2i}.$$

Also, one has

$$|z|^2 = z\overline{z} \quad \text{and as a consequence} \quad \frac{1}{z} = \frac{\overline{z}}{|z|^2} \text{ whenever } z \neq 0.$$

Any non-zero complex number z can be written in **polar form**

$$z = re^{i\theta},$$

where $r > 0$; also $\theta \in \mathbb{R}$ is called the **argument** of z (defined uniquely up to a multiple of 2π) and is often denoted by $\arg z$, and

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Since $|e^{i\theta}| = 1$ we observe that $r = |z|$, and θ is simply the angle (with positive counterclockwise orientation) between the positive real axis and the half-line starting at the origin and passing through z. (See Figure 2.)

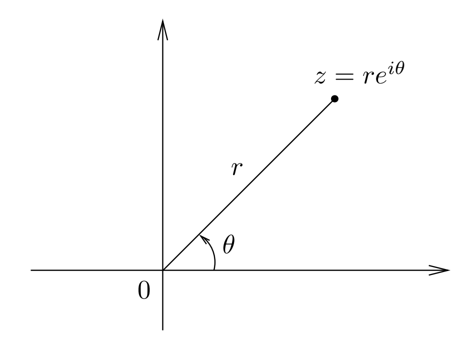


Figure 2. The polar form of a complex number

Finally, note that if $z = re^{i\theta}$ and $w = se^{i\varphi}$, then

$$zw = rse^{i(\theta+\varphi)},$$

so multiplication by a complex number corresponds to a homothety in $\mathbb{R}^2$ (that is, a rotation composed with a dilation).

1.2 Convergence

We make a transition from the arithmetic and geometric properties of complex numbers described above to the key notions of convergence and limits.

A sequence $\{z_1, z_2, \ldots\}$ of complex numbers is said to **converge** to $w \in \mathbb{C}$ if

$$\lim_{n\to\infty} |z_n - w| = 0, \quad \text{and we write} \quad w = \lim_{n\to\infty} z_n.$$

This notion of convergence is not new. Indeed, since absolute values in $\mathbb{C}$ and Euclidean distances in $\mathbb{R}^2$ coincide, we see that z_n converges to w if and only if the corresponding sequence of points in the complex plane converges to the point that corresponds to w.

As an exercise, the reader can check that the sequence $\{z_n\}$ converges to w if and only if the sequence of real and imaginary parts of z_n converge to the real and imaginary parts of w, respectively.

Since it is sometimes not possible to readily identify the limit of a sequence (for example, $\lim_{N\to\infty} \sum_{n=1}^{N} 1/n^3$), it is convenient to have a condition on the sequence itself which is equivalent to its convergence. A sequence $\{z_n\}$ is said to be a **Cauchy sequence** (or simply **Cauchy**) if

$$|z_n - z_m| \to 0 \quad \text{as } n, m \to \infty.$$

In other words, given $\epsilon > 0$ there exists an integer $N > 0$ so that $|z_n - z_m| < \epsilon$ whenever $n, m > N$. An important fact of real analysis is that $\mathbb{R}$ is complete: every Cauchy sequence of real numbers converges to a real number.[1] Since the sequence $\{z_n\}$ is Cauchy if and only if the sequences of real and imaginary parts of z_n are, we conclude that every Cauchy sequence in $\mathbb{C}$ has a limit in $\mathbb{C}$. We have thus the following result.

Theorem 1.1 $\mathbb{C}$, *the complex numbers, is complete.*

We now turn our attention to some simple topological considerations that are necessary in our study of functions. Here again, the reader will note that no new notions are introduced, but rather previous notions are now presented in terms of a new vocabulary.

1.3 Sets in the complex plane

If $z_0 \in \mathbb{C}$ and $r > 0$, we define the **open disc $D_r(z_0)$ of radius r centered at z_0** to be the set of all complex numbers that are at absolute

---

[1] This is sometimes called the Bolzano-Weierstrass theorem.

value strictly less than r from z_0. In other words,

$$D_r(z_0) = \{z \in \mathbb{C} : |z - z_0| < r\},$$

and this is precisely the usual disc in the plane of radius r centered at z_0. The **closed disc $\overline{D}_r(z_0)$ of radius r centered at** z_0 is defined by

$$\overline{D}_r(z_0) = \{z \in \mathbb{C} : |z - z_0| \leq r\},$$

and the boundary of either the open or closed disc is the circle

$$C_r(z_0) = \{z \in \mathbb{C} : |z - z_0| = r\}.$$

Since the **unit disc** (that is, the open disc centered at the origin and of radius 1) plays an important role in later chapters, we will often denote it by $\mathbb{D}$,

$$\mathbb{D} = \{z \in \mathbb{C} : \ |z| < 1\}.$$

Given a set $\Omega \subset \mathbb{C}$, a point z_0 is an **interior point** of Ω if there exists $r > 0$ such that

$$D_r(z_0) \subset \Omega.$$

The **interior** of Ω consists of all its interior points. Finally, a set Ω is **open** if every point in that set is an interior point of Ω. This definition coincides precisely with the definition of an open set in $\mathbb{R}^2$.

A set Ω is **closed** if its complement $\Omega^c = \mathbb{C} - \Omega$ is open. This property can be reformulated in terms of limit points. A point $z \in \mathbb{C}$ is said to be a **limit point** of the set Ω if there exists a sequence of points $z_n \in \Omega$ such that $z_n \neq z$ and $\lim_{n\to\infty} z_n = z$. The reader can now check that a set is closed if and only if it contains all its limit points. The **closure** of any set Ω is the union of Ω and its limit points, and is often denoted by $\overline{\Omega}$.

Finally, the **boundary** of a set Ω is equal to its closure minus its interior, and is often denoted by $\partial\Omega$.

A set Ω is **bounded** if there exists $M > 0$ such that $|z| < M$ whenever $z \in \Omega$. In other words, the set Ω is contained in some large disc. If Ω is bounded, we define its **diameter** by

$$\operatorname{diam}(\Omega) = \sup_{z,w\in\Omega} |z - w|.$$

A set Ω is said to be **compact** if it is closed and bounded. Arguing as in the case of real variables, one can prove the following.

Theorem 1.2 *The set $\Omega \subset \mathbb{C}$ is compact if and only if every sequence $\{z_n\} \subset \Omega$ has a subsequence that converges to a point in Ω.*

An **open covering** of Ω is a family of open sets $\{U_\alpha\}$ (not necessarily countable) such that

$$\Omega \subset \bigcup_{\alpha} U_\alpha.$$

In analogy with the situation in $\mathbb{R}$, we have the following equivalent formulation of compactness.

Theorem 1.3 *A set Ω is compact if and only if every open covering of Ω has a finite subcovering.*

Another interesting property of compactness is that of **nested sets**. We shall in fact use this result at the very beginning of our study of complex function theory, more precisely in the proof of Goursat's theorem in Chapter 2.

Proposition 1.4 *If $\Omega_1 \supset \Omega_2 \supset \cdots \supset \Omega_n \supset \cdots$ is a sequence of non-empty compact sets in $\mathbb{C}$ with the property that*

$$\text{diam}(\Omega_n) \to 0 \qquad \textit{as } n \to \infty,$$

then there exists a unique point $w \in \mathbb{C}$ such that $w \in \Omega_n$ for all n.

Proof. Choose a point z_n in each Ω_n. The condition $\text{diam}(\Omega_n) \to 0$ says precisely that $\{z_n\}$ is a Cauchy sequence, therefore this sequence converges to a limit that we call w. Since each set Ω_n is compact we must have $w \in \Omega_n$ for all n. Finally, w is the unique point satisfying this property, for otherwise, if w' satisfied the same property with $w' \neq w$ we would have $|w - w'| > 0$ and the condition $\text{diam}(\Omega_n) \to 0$ would be violated.

The last notion we need is that of connectedness. An open set $\Omega \subset \mathbb{C}$ is said to be **connected** if it is not possible to find two disjoint non-empty open sets Ω_1 and Ω_2 such that

$$\Omega = \Omega_1 \cup \Omega_2.$$

A connected open set in $\mathbb{C}$ will be called a **region**. Similarly, a closed set F is connected if one cannot write $F = F_1 \cup F_2$ where F_1 and F_2 are disjoint non-empty closed sets.

There is an equivalent definition of connectedness for open sets in terms of curves, which is often useful in practice: an open set Ω is connected if and only if any two points in Ω can be joined by a curve γ entirely contained in Ω. See Exercise 5 for more details.

2 Functions on the complex plane

2.1 Continuous functions

Let f be a function defined on a set Ω of complex numbers. We say that f is **continuous** at the point $z_0 \in \Omega$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that whenever $z \in \Omega$ and $|z - z_0| < \delta$ then $|f(z) - f(z_0)| < \epsilon$. An equivalent definition is that for every sequence $\{z_1, z_2, \ldots\} \subset \Omega$ such that $\lim z_n = z_0$, then $\lim f(z_n) = f(z_0)$.

The function f is said to be continuous on Ω if it is continuous at every point of Ω. Sums and products of continuous functions are also continuous.

Since the notions of convergence for complex numbers and points in $\mathbb{R}^2$ are the same, the function f of the complex argument $z = x + iy$ is continuous if and only if it is continuous viewed as a function of the two real variables x and y.

By the triangle inequality, it is immediate that if f is continuous, then the real-valued function defined by $z \mapsto |f(z)|$ is continuous. We say that f attains a **maximum** at the point $z_0 \in \Omega$ if

$$|f(z)| \leq |f(z_0)| \quad \text{for all } z \in \Omega,$$

with the inequality reversed for the definition of a **minimum**.

Theorem 2.1 *A continuous function on a compact set Ω is bounded and attains a maximum and minimum on Ω.*

This is of course analogous to the situation of functions of a real variable, and we shall not repeat the simple proof here.

2.2 Holomorphic functions

We now present a notion that is central to complex analysis, and in distinction to our previous discussion we introduce a definition that is genuinely *complex* in nature.

Let Ω be an open set in $\mathbb{C}$ and f a complex-valued function on Ω. The function f is **holomorphic at the point** $z_0 \in \Omega$ if the quotient

$$\frac{f(z_0 + h) - f(z_0)}{h} \tag{1}$$

converges to a limit when $h \to 0$. Here $h \in \mathbb{C}$ and $h \neq 0$ with $z_0 + h \in \Omega$, so that the quotient is well defined. The limit of the quotient, when it exists, is denoted by $f'(z_0)$, and is called the **derivative of f at z_0**:

$$f'(z_0) = \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}.$$

It should be emphasized that in the above limit, h is a complex number that may approach 0 from any direction.

The function f is said to be **holomorphic on** Ω if f is holomorphic at every point of Ω. If C is a closed subset of $\mathbb{C}$, we say that f is **holomorphic on** C if f is holomorphic in some open set containing C. Finally, if f is holomorphic in all of $\mathbb{C}$ we say that f is **entire**.

Sometimes the terms **regular** or **complex differentiable** are used instead of holomorphic. The latter is natural in view of (1) which mimics the usual definition of the derivative of a function of one real variable. But despite this resemblance, a holomorphic function of one complex variable will satisfy much stronger properties than a differentiable function of one real variable. For example, a holomorphic function will actually be infinitely many times complex differentiable, that is, the existence of the first derivative will guarantee the existence of derivatives of any order. This is in contrast with functions of one real variable, since there are differentiable functions that do not have two derivatives. In fact more is true: every holomorphic function is analytic, in the sense that it has a power series expansion near every point (power series will be discussed in the next section), and for this reason we also use the term **analytic** as a synonym for holomorphic. Again, this is in contrast with the fact that there are indefinitely differentiable functions of one real variable that cannot be expanded in a power series. (See Exercise 23.)

Example 1. The function $f(z) = z$ is holomorphic on any open set in $\mathbb{C}$, and $f'(z) = 1$. In fact, any polynomial

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n$$

is holomorphic in the entire complex plane and

$$p'(z) = a_1 + \cdots + n a_n z^{n-1}.$$

This follows from Proposition 2.2 below.

Example 2. The function $1/z$ is holomorphic on any open set in $\mathbb{C}$ that does not contain the origin, and $f'(z) = -1/z^2$.

Example 3. The function $f(z) = \overline{z}$ is not holomorphic. Indeed, we have

$$\frac{f(z_0 + h) - f(z_0)}{h} = \frac{\overline{h}}{h}$$

which has no limit as $h \to 0$, as one can see by first taking h real and then h purely imaginary.

An important family of examples of holomorphic functions, which we discuss later, are the power series. They contain functions such as $e^z, \sin z$, or $\cos z$, and in fact power series play a crucial role in the theory of holomorphic functions, as we already mentioned in the last paragraph. Some other examples of holomorphic functions that will make their appearance in later chapters were given in the introduction to this book.

It is clear from (1) above that a function f is holomorphic at $z_0 \in \Omega$ if and only if there exists a complex number a such that

$$f(z_0 + h) - f(z_0) - ah = h\psi(h), \tag{2}$$

where ψ is a function defined for all small h and $\lim_{h\to 0} \psi(h) = 0$. Of course, we have $a = f'(z_0)$. From this formulation, it is clear that f is continuous wherever it is holomorphic. Arguing as in the case of one real variable, using formulation (2) in the case of the chain rule (for example), one proves easily the following desirable properties of holomorphic functions.

Proposition 2.2 *If f and g are holomorphic in Ω, then:*

(i) *$f + g$ is holomorphic in Ω and $(f + g)' = f' + g'$.*

(ii) *fg is holomorphic in Ω and $(fg)' = f'g + fg'$.*

(iii) *If $g(z_0) \neq 0$, then f/g is holomorphic at z_0 and*

$$(f/g)' = \frac{f'g - fg'}{g^2}.$$

Moreover, if $f : \Omega \to U$ and $g : U \to \mathbb{C}$ are holomorphic, the chain rule holds

$$(g \circ f)'(z) = g'(f(z))f'(z) \quad \textit{for all } z \in \Omega.$$

Complex-valued functions as mappings

We now clarify the relationship between the complex and real derivatives. In fact, the third example above should convince the reader that the notion of complex differentiability differs significantly from the usual notion of real differentiability of a function of two real variables. Indeed, in terms of real variables, the function $f(z) = \overline{z}$ corresponds to the map $F : (x, y) \mapsto (x, -y)$, which is differentiable in the real sense. Its derivative at a point is the linear map given by its Jacobian, the 2×2 matrix of partial derivatives of the coordinate functions. In fact, F is linear and

is therefore equal to its derivative at every point. This implies that F is actually indefinitely differentiable. In particular the existence of the real derivative need not guarantee that f is holomorphic.

This example leads us to associate more generally to each complex-valued function $f = u + iv$ the mapping $F(x, y) = (u(x, y), v(x, y))$ from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Recall that a function $F(x, y) = (u(x, y), v(x, y))$ is said to be differentiable at a point $P_0 = (x_0, y_0)$ if there exists a linear transformation $J : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$\frac{|F(P_0 + H) - F(P_0) - J(H)|}{|H|} \to 0 \quad \text{as } |H| \to 0,\ H \in \mathbb{R}^2. \tag{3}$$

Equivalently, we can write

$$F(P_0 + H) - F(P_0) = J(H) + |H|\Psi(H),$$

with $|\Psi(H)| \to 0$ as $|H| \to 0$. The linear transformation J is unique and is called the derivative of F at P_0. If F is differentiable, the partial derivatives of u and v exist, and the linear transformation J is described in the standard basis of $\mathbb{R}^2$ by the Jacobian matrix of F

$$J = J_F(x, y) = \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix}.$$

In the case of complex differentiation the derivative is a complex number $f'(z_0)$, while in the case of real derivatives, it is a matrix. There is, however, a connection between these two notions, which is given in terms of special relations that are satisfied by the entries of the Jacobian matrix, that is, the partials of u and v. To find these relations, consider the limit in (1) when h is first real, say $h = h_1 + ih_2$ with $h_2 = 0$. Then, if we write $z = x + iy$, $z_0 = x_0 + iy_0$, and $f(z) = f(x, y)$, we find that

$$\begin{aligned} f'(z_0) &= \lim_{h_1 \to 0} \frac{f(x_0 + h_1, y_0) - f(x_0, y_0)}{h_1} \\ &= \frac{\partial f}{\partial x}(z_0), \end{aligned}$$

where $\partial/\partial x$ denotes the usual partial derivative in the x variable. (We fix y_0 and think of f as a complex-valued function of the single real variable x.) Now taking h purely imaginary, say $h = ih_2$, a similar argument yields

$$
\begin{aligned}
f'(z_0) &= \lim_{h_2 \to 0} \frac{f(x_0, y_0 + h_2) - f(x_0, y_0)}{ih_2} \\
&= \frac{1}{i} \frac{\partial f}{\partial y}(z_0),
\end{aligned}
$$

where $\partial/\partial y$ is partial differentiation in the y variable. Therefore, if f is holomorphic we have shown that

$$
\frac{\partial f}{\partial x} = \frac{1}{i} \frac{\partial f}{\partial y}.
$$

Writing $f = u + iv$, we find after separating real and imaginary parts and using $1/i = -i$, that the partials of u and v exist, and they satisfy the following non-trivial relations

$$
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.
$$

These are the **Cauchy-Riemann** equations, which link real and complex analysis.

We can clarify the situation further by defining two differential operators

$$
\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} + \frac{1}{i}\frac{\partial}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{1}{i}\frac{\partial}{\partial y}\right).
$$

Proposition 2.3 *If f is holomorphic at z_0, then*

$$
\frac{\partial f}{\partial \overline{z}}(z_0) = 0 \quad \textit{and} \quad f'(z_0) = \frac{\partial f}{\partial z}(z_0) = 2\frac{\partial u}{\partial z}(z_0).
$$

Also, if we write $F(x, y) = f(z)$, then F is differentiable in the sense of real variables, and

$$
\det J_F(x_0, y_0) = |f'(z_0)|^2.
$$

Proof. Taking real and imaginary parts, it is easy to see that the Cauchy-Riemann equations are equivalent to $\partial f/\partial \overline{z} = 0$. Moreover, by our earlier observation

$$
\begin{aligned}
f'(z_0) &= \frac{1}{2}\left(\frac{\partial f}{\partial x}(z_0) + \frac{1}{i}\frac{\partial f}{\partial y}(z_0)\right) \\
&= \frac{\partial f}{\partial z}(z_0),
\end{aligned}
$$

and the Cauchy-Riemann equations give $\partial f/\partial z = 2\partial u/\partial z$. To prove that F is differentiable it suffices to observe that if $H = (h_1, h_2)$ and $h = h_1 + ih_2$, then the Cauchy-Riemann equations imply

$$J_F(x_0, y_0)(H) = \left(\frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}\right)(h_1 + ih_2) = f'(z_0)h\,,$$

where we have identified a complex number with the pair of real and imaginary parts. After a final application of the Cauchy-Riemann equations, the above results imply that

$$\det J_F(x_0, y_0) = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y} = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = \left|2\frac{\partial u}{\partial z}\right|^2 = |f'(z_0)|^2. \tag{4}$$

So far, we have assumed that f is holomorphic and deduced relations satisfied by its real and imaginary parts. The next theorem contains an important converse, which completes the circle of ideas presented here.

Theorem 2.4 *Suppose $f = u + iv$ is a complex-valued function defined on an open set Ω. If u and v are continuously differentiable and satisfy the Cauchy-Riemann equations on Ω, then f is holomorphic on Ω and $f'(z) = \partial f/\partial z$.*

Proof. Write

$$u(x + h_1, y + h_2) - u(x, y) = \frac{\partial u}{\partial x}h_1 + \frac{\partial u}{\partial y}h_2 + |h|\psi_1(h)$$

and

$$v(x + h_1, y + h_2) - v(x, y) = \frac{\partial v}{\partial x}h_1 + \frac{\partial v}{\partial y}h_2 + |h|\psi_2(h),$$

where $\psi_j(h) \to 0$ (for $j = 1, 2$) as $|h|$ tends to 0, and $h = h_1 + ih_2$. Using the Cauchy-Riemann equations we find that

$$f(z + h) - f(z) = \left(\frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}\right)(h_1 + ih_2) + |h|\psi(h),$$

where $\psi(h) = \psi_1(h) + \psi_2(h) \to 0$, as $|h| \to 0$. Therefore f is holomorphic and

$$f'(z) = 2\frac{\partial u}{\partial z} = \frac{\partial f}{\partial z}.$$

2.3 Power series

The prime example of a power series is the complex **exponential** function, which is defined for $z \in \mathbb{C}$ by

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

When z is real, this definition coincides with the usual exponential function, and in fact, the series above converges absolutely for every $z \in \mathbb{C}$. To see this, note that

$$\left| \frac{z^n}{n!} \right| = \frac{|z|^n}{n!},$$

so $|e^z|$ can be compared to the series $\sum |z|^n/n! = e^{|z|} < \infty$. In fact, this estimate shows that the series defining e^z is uniformly convergent in every disc in $\mathbb{C}$.

In this section we will prove that e^z is holomorphic in all of $\mathbb{C}$ (it is entire), and that its derivative can be found by differentiating the series term by term. Hence

$$(e^z)' = \sum_{n=0}^{\infty} n \frac{z^{n-1}}{n!} = \sum_{m=0}^{\infty} \frac{z^m}{m!} = e^z,$$

and therefore e^z is its own derivative.

In contrast, the geometric series

$$\sum_{n=0}^{\infty} z^n$$

converges absolutely only in the disc $|z| < 1$, and its sum there is the function $1/(1-z)$, which is holomorphic in the open set $\mathbb{C} - \{1\}$. This identity is proved exactly as when z is real: we first observe

$$\sum_{n=0}^{N} z^n = \frac{1 - z^{N+1}}{1 - z},$$

and then note that if $|z| < 1$ we must have $\lim_{N \to \infty} z^{N+1} = 0$.

In general, a **power series** is an expansion of the form

$$\sum_{n=0}^{\infty} a_n z^n, \tag{5}$$

where $a_n \in \mathbb{C}$. To test for absolute convergence of this series, we must investigate

$$\sum_{n=0}^{\infty} |a_n| \, |z|^n ,$$

and we observe that if the series (5) converges absolutely for some z_0, then it will also converge for all z in the disc $|z| \leq |z_0|$. We now prove that there always exists an open disc (possibly empty) on which the power series converges absolutely.

Theorem 2.5 *Given a power series* $\sum_{n=0}^{\infty} a_n z^n$*, there exists* $0 \leq R \leq \infty$ *such that:*

(i) *If* $|z| < R$ *the series converges absolutely.*

(ii) *If* $|z| > R$ *the series diverges.*

Moreover, if we use the convention that $1/0 = \infty$ *and* $1/\infty = 0$*, then* R *is given by Hadamard's formula*

$$1/R = \limsup \, |a_n|^{1/n}.$$

The number R is called the **radius of convergence** of the power series, and the region $|z| < R$ the **disc of convergence**. In particular, we have $R = \infty$ in the case of the exponential function, and $R = 1$ for the geometric series.

Proof. Let $L = 1/R$ where R is defined by the formula in the statement of the theorem, and suppose that $L \neq 0, \infty$. (These two easy cases are left as an exercise.) If $|z| < R$, choose $\epsilon > 0$ so small that

$$(L + \epsilon)|z| = r < 1.$$

By the definition L, we have $|a_n|^{1/n} \leq L + \epsilon$ for all large n, therefore

$$|a_n| \, |z|^n \leq \{(L + \epsilon)|z|\}^n = r^n.$$

Comparison with the geometric series $\sum r^n$ shows that $\sum a_n z^n$ converges.

If $|z| > R$, then a similar argument proves that there exists a sequence of terms in the series whose absolute value goes to infinity, hence the series diverges.

Remark. On the boundary of the disc of convergence, $|z| = R$, the situation is more delicate as one can have either convergence or divergence. (See Exercise 19.)

Further examples of power series that converge in the whole complex plane are given by the standard **trigonometric functions**; these are defined by

$$\cos z = \sum_{n=0}^{\infty}(-1)^n \frac{z^{2n}}{(2n)!}, \qquad \text{and} \qquad \sin z = \sum_{n=0}^{\infty}(-1)^n \frac{z^{2n+1}}{(2n+1)!},$$

and they agree with the usual cosine and sine of a real argument whenever $z \in \mathbb{R}$. A simple calculation exhibits a connection between these two functions and the complex exponential, namely,

$$\cos z = \frac{e^{iz}+e^{-iz}}{2} \qquad \text{and} \qquad \sin z = \frac{e^{iz}-e^{-iz}}{2i}.$$

These are called the **Euler formulas** for the cosine and sine functions.

Power series provide a very important class of analytic functions that are particularly simple to manipulate.

Theorem 2.6 *The power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ defines a holomorphic function in its disc of convergence. The derivative of f is also a power series obtained by differentiating term by term the series for f, that is,*

$$f'(z) = \sum_{n=0}^{\infty} n a_n z^{n-1}.$$

Moreover, f' has the same radius of convergence as f.

Proof. The assertion about the radius of convergence of f' follows from Hadamard's formula. Indeed, $\lim_{n\to\infty} n^{1/n} = 1$, and therefore

$$\limsup |a_n|^{1/n} = \limsup |na_n|^{1/n},$$

so that $\sum a_n z^n$ and $\sum n a_n z^n$ have the same radius of convergence, and hence so do $\sum a_n z^n$ and $\sum n a_n z^{n-1}$.

To prove the first assertion, we must show that the series

$$g(z) = \sum_{n=0}^{\infty} n a_n z^{n-1}$$

gives the derivative of f. For that, let R denote the radius of convergence of f, and suppose $|z_0| < r < R$. Write

$$f(z) = S_N(z) + E_N(z)\,,$$

where

$$S_N(z) = \sum_{n=0}^{N} a_n z^n \quad \text{and} \quad E_N(z) = \sum_{n=N+1}^{\infty} a_n z^n.$$

Then, if h is chosen so that $|z_0 + h| < r$ we have

$$\frac{f(z_0+h)-f(z_0)}{h} - g(z_0) = \left(\frac{S_N(z_0+h)-S_N(z_0)}{h} - S_N'(z_0)\right) + (S_N'(z_0) - g(z_0)) + \left(\frac{E_N(z_0+h)-E_N(z_0)}{h}\right).$$

Since $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$, we see that

$$\left|\frac{E_N(z_0+h)-E_N(z_0)}{h}\right| \leq \sum_{n=N+1}^{\infty} |a_n| \left|\frac{(z_0+h)^n - z_0^n}{h}\right| \leq \sum_{n=N+1}^{\infty} |a_n| n r^{n-1},$$

where we have used the fact that $|z_0| < r$ and $|z_0 + h| < r$. The expression on the right is the tail end of a convergent series, since g converges absolutely on $|z| < R$. Therefore, given $\epsilon > 0$ we can find N_1 so that $N > N_1$ implies

$$\left|\frac{E_N(z_0+h)-E_N(z_0)}{h}\right| < \epsilon.$$

Also, since $\lim_{N\to\infty} S_N'(z_0) = g(z_0)$, we can find N_2 so that $N > N_2$ implies

$$|S_N'(z_0) - g(z_0)| < \epsilon.$$

If we fix N so that both $N > N_1$ and $N > N_2$ hold, then we can find $\delta > 0$ so that $|h| < \delta$ implies

$$\left|\frac{S_N(z_0+h)-S_N(z_0)}{h} - S_N'(z_0)\right| < \epsilon\,,$$

simply because the derivative of a polynomial is obtained by differentiating it term by term. Therefore,

$$\left|\frac{f(z_0+h)-f(z_0)}{h} - g(z_0)\right| < 3\epsilon$$

whenever $|h| < \delta$, thereby concluding the proof of the theorem.

Successive applications of this theorem yield the following.

Corollary 2.7 *A power series is infinitely complex differentiable in its disc of convergence, and the higher derivatives are also power series obtained by termwise differentiation.*

We have so far dealt only with power series centered at the origin. More generally, a power series centered at $z_0 \in \mathbb{C}$ is an expression of the form

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

The disc of convergence of f is now centered at z_0 and its radius is still given by Hadamard's formula. In fact, if

$$g(z) = \sum_{n=0}^{\infty} a_n z^n,$$

then f is simply obtained by translating g, namely $f(z) = g(w)$ where $w = z - z_0$. As a consequence everything about g also holds for f after we make the appropriate translation. In particular, by the chain rule,

$$f'(z) = g'(w) = \sum_{n=0}^{\infty} n a_n (z - z_0)^{n-1}.$$

A function f defined on an open set Ω is said to be **analytic** (or have a **power series expansion**) at a point $z_0 \in \Omega$ if there exists a power series $\sum a_n (z - z_0)^n$ centered at z_0, with positive radius of convergence, such that

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad \text{for all } z \text{ in a neighborhood of } z_0.$$

If f has a power series expansion at every point in Ω, we say that f is **analytic on** Ω.

By Theorem 2.6, an analytic function on Ω is also holomorphic there. A deep theorem which we prove in the next chapter says that the converse is true: every holomorphic function is analytic. For that reason, we use the terms holomorphic and analytic interchangeably.

3 Integration along curves

In the definition of a curve, we distinguish between the one-dimensional geometric object in the plane (endowed with an orientation), and its

parametrization, which is a mapping from a closed interval to $\mathbb{C}$, that is not uniquely determined.

A **parametrized curve** is a function $z(t)$ which maps a closed interval $[a, b] \subset \mathbb{R}$ to the complex plane. We shall impose regularity conditions on the parametrization which are always verified in the situations that occur in this book. We say that the parametrized curve is **smooth** if $z'(t)$ exists and is continuous on $[a, b]$, and $z'(t) \neq 0$ for $t \in [a, b]$. At the points $t = a$ and $t = b$, the quantities $z'(a)$ and $z'(b)$ are interpreted as the one-sided limits

$$z'(a) = \lim_{\substack{h \to 0 \\ h > 0}} \frac{z(a+h) - z(a)}{h} \quad \text{and} \quad z'(b) = \lim_{\substack{h \to 0 \\ h < 0}} \frac{z(b+h) - z(b)}{h}.$$

In general, these quantities are called the right-hand derivative of $z(t)$ at a, and the left-hand derivative of $z(t)$ at b, respectively.

Similarly we say that the parametrized curve is **piecewise-smooth** if z is continuous on $[a, b]$ and if there exist points

$$a = a_0 < a_1 < \cdots < a_n = b\,,$$

where $z(t)$ is smooth in the intervals $[a_k, a_{k+1}]$. In particular, the right-hand derivative at a_k may differ from the left-hand derivative at a_k for $k = 1, \ldots, n-1$.

Two parametrizations,

$$z : [a, b] \to \mathbb{C} \quad \text{and} \quad \tilde{z} : [c, d] \to \mathbb{C},$$

are **equivalent** if there exists a continuously differentiable bijection $s \mapsto t(s)$ from $[c, d]$ to $[a, b]$ so that $t'(s) > 0$ and

$$\tilde{z}(s) = z(t(s)).$$

The condition $t'(s) > 0$ says precisely that the orientation is preserved: as s travels from c to d, then $t(s)$ travels from a to b. The family of all parametrizations that are equivalent to $z(t)$ determines a **smooth curve** $\gamma \subset \mathbb{C}$, namely the image of $[a, b]$ under z with the orientation given by z as t travels from a to b. We can define a curve γ^- obtained from the curve γ by reversing the orientation (so that γ and γ^- consist of the same points in the plane). As a particular parametrization for γ^- we can take $z^- : [a, b] \to \mathbb{R}^2$ defined by

$$z^-(t) = z(b + a - t).$$

It is also clear how to define a **piecewise-smooth curve**. The points $z(a)$ and $z(b)$ are called the **end-points** of the curve and are independent on the parametrization. Since γ carries an orientation, it is natural to say that γ begins at $z(a)$ and ends at $z(b)$.

A smooth or piecewise-smooth curve is **closed** if $z(a) = z(b)$ for any of its parametrizations. Finally, a smooth or piecewise-smooth curve is **simple** if it is not self-intersecting, that is, $z(t) \neq z(s)$ unless $s = t$. Of course, if the curve is closed to begin with, then we say that it is simple whenever $z(t) \neq z(s)$ unless $s = t$, or $s = a$ and $t = b$.

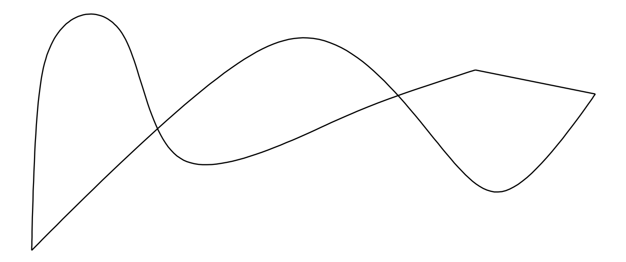


Figure 3. A closed piecewise-smooth curve

For brevity, we shall call any piecewise-smooth curve a **curve**, since these will be the objects we shall be primarily concerned with.

A basic example consists of a circle. Consider the circle $C_r(z_0)$ centered at z_0 and of radius r, which by definition is the set

$$C_r(z_0) = \{z \in \mathbb{C} : \ |z - z_0| = r\}.$$

The **positive orientation** (counterclockwise) is the one that is given by the standard parametrization

$$z(t) = z_0 + re^{it}, \qquad \text{where } t \in [0, 2\pi],$$

while the **negative orientation** (clockwise) is given by

$$z(t) = z_0 + re^{-it}, \qquad \text{where } t \in [0, 2\pi].$$

In the following chapters, we shall denote by C a general *positively* oriented circle.

An important tool in the study of holomorphic functions is integration of functions along curves. Loosely speaking, a key theorem in complex

analysis says that if a function is holomorphic in the interior of a closed curve γ, then

$$\int_\gamma f(z)\,dz = 0,$$

and we shall turn our attention to a version of this theorem (called Cauchy's theorem) in the next chapter. Here we content ourselves with the necessary definitions and properties of the integral.

Given a smooth curve γ in $\mathbb{C}$ parametrized by $z : [a, b] \to \mathbb{C}$, and f a continuous function on γ, we define the **integral of f along γ** by

$$\int_\gamma f(z)\,dz = \int_a^b f(z(t))z'(t)\,dt.$$

In order for this definition to be meaningful, we must show that the right-hand integral is independent of the parametrization chosen for γ. Say that $\tilde{z}$ is an equivalent parametrization as above. Then the change of variables formula and the chain rule imply that

$$\int_a^b f(z(t))z'(t)\,dt = \int_c^d f(z(t(s)))z'(t(s))t'(s)\,ds = \int_c^d f(\tilde{z}(s))\tilde{z}'(s)\,ds.$$

This proves that the integral of f over γ is well defined.

If γ is piecewise smooth, then the integral of f over γ is simply the sum of the integrals of f over the smooth parts of γ, so if $z(t)$ is a piecewise-smooth parametrization as before, then

$$\int_\gamma f(z)\,dz = \sum_{k=0}^{n-1} \int_{a_k}^{a_{k+1}} f(z(t))z'(t)\,dt.$$

By definition, the **length** of the smooth curve γ is

$$\text{length}(\gamma) = \int_a^b |z'(t)|\,dt.$$

Arguing as we just did, it is clear that this definition is also independent of the parametrization. Also, if γ is only piecewise-smooth, then its length is the sum of the lengths of its smooth parts.

Proposition 3.1 *Integration of continuous functions over curves satisfies the following properties:*

(i) *It is linear, that is, if* $\alpha, \beta \in \mathbb{C}$*, then*

$$\int_\gamma (\alpha f(z) + \beta g(z))\, dz = \alpha \int_\gamma f(z)\, dz + \beta \int_\gamma g(z)\, dz.$$

(ii) *If* γ^- *is* γ *with the reverse orientation, then*

$$\int_\gamma f(z)\, dz = -\int_{\gamma^-} f(z)\, dz.$$

(iii) *One has the inequality*

$$\left| \int_\gamma f(z)\, dz \right| \le \sup_{z \in \gamma} |f(z)| \cdot \text{length}(\gamma).$$

Proof. The first property follows from the definition and the linearity of the Riemann integral. The second property is left as an exercise. For the third, note that

$$\left| \int_\gamma f(z)\, dz \right| \le \sup_{t \in [a,b]} |f(z(t))| \int_a^b |z'(t)|\, dt \le \sup_{z \in \gamma} |f(z)| \cdot \text{length}(\gamma)$$

as was to be shown.

As we have said, Cauchy's theorem states that for appropriate closed curves γ in an open set Ω on which f is holomorphic, then

$$\int_\gamma f(z)\, dz = 0.$$

The existence of primitives gives a first manifestation of this phenomenon. Suppose f is a function on the open set Ω. A **primitive** for f on Ω is a function F that is holomorphic on Ω and such that $F'(z) = f(z)$ for all $z \in \Omega$.

Theorem 3.2 *If a continuous function* f *has a primitive* F *in* Ω*, and* γ *is a curve in* Ω *that begins at* w_1 *and ends at* w_2*, then*

$$\int_\gamma f(z)\, dz = F(w_2) - F(w_1).$$

Proof. If γ is smooth, the proof is a simple application of the chain rule and the fundamental theorem of calculus. Indeed, if $z(t) : [a, b] \to \mathbb{C}$ is a parametrization for γ, then $z(a) = w_1$ and $z(b) = w_2$, and we have

$$\begin{aligned}\int_\gamma f(z)\,dz &= \int_a^b f(z(t))z'(t)\,dt\\ &= \int_a^b F'(z(t))z'(t)\,dt\\ &= \int_a^b \frac{d}{dt}F(z(t))\,dt\\ &= F(z(b)) - F(z(a)).\end{aligned}$$

If γ is only piecewise-smooth, then arguing as we just did, we obtain a telescopic sum, and we have

$$\begin{aligned}\int_\gamma f(z)\,dz &= \sum_{k=0}^{n-1} F(z(a_{k+1})) - F(z(a_k))\\ &= F(z(a_n)) - F(z(a_0))\\ &= F(z(b)) - F(z(a)).\end{aligned}$$

Corollary 3.3 *If γ is a closed curve in an open set Ω, and f is continuous and has a primitive in Ω, then*

$$\int_\gamma f(z)\,dz = 0.$$

This is immediate since the end-points of a closed curve coincide.

For example, the function $f(z) = 1/z$ does not have a primitive in the open set $\mathbb{C} - \{0\}$, since if C is the unit circle parametrized by $z(t) = e^{it}$, $0 \le t \le 2\pi$, we have

$$\int_C f(z)\,dz = \int_0^{2\pi} \frac{ie^{it}}{e^{it}}\,dt = 2\pi i \neq 0.$$

In subsequent chapters, we shall see that this innocent calculation, which provides an example of a function f and closed curve γ for which $\int_\gamma f(z)\,dz \neq 0$, lies at the heart of the theory.

Corollary 3.4 *If f is holomorphic in a region Ω and $f' = 0$, then f is constant.*

Proof. Fix a point $w_0 \in \Omega$. It suffices to show that $f(w) = f(w_0)$ for all $w \in \Omega$.

Since Ω is connected, for any $w \in \Omega$, there exists a curve γ which joins w_0 to w. Since f is clearly a primitive for f', we have

$$\int_\gamma f'(z)\, dz = f(w) - f(w_0).$$

By assumption, $f' = 0$ so the integral on the left is 0, and we conclude that $f(w) = f(w_0)$ as desired.

Remark on notation. When convenient, we follow the practice of using the notation $f(z) = O(g(z))$ to mean that there is a constant $C > 0$ such that $|f(z)| \leq C|g(z)|$ for z in a neighborhood of the point in question. In addition, we say $f(z) = o(g(z))$ when $|f(z)/g(z)| \to 0$. We also write $f(z) \sim g(z)$ to mean that $f(z)/g(z) \to 1$.

4 Exercises

1. Describe geometrically the sets of points z in the complex plane defined by the following relations:

(a) $|z - z_1| = |z - z_2|$ where $z_1, z_2 \in \mathbb{C}$.

(b) $1/z = \overline{z}$.

(c) $\text{Re}(z) = 3$.

(d) $\text{Re}(z) > c$, (resp., $\geq c$) where $c \in \mathbb{R}$.

(e) $\text{Re}(az + b) > 0$ where $a, b \in \mathbb{C}$.

(f) $|z| = \text{Re}(z) + 1$.

(g) $\text{Im}(z) = c$ with $c \in \mathbb{R}$.

2. Let $\langle \cdot, \cdot \rangle$ denote the usual inner product in $\mathbb{R}^2$. In other words, if $Z = (x_1, y_1)$ and $W = (x_2, y_2)$, then

$$\langle Z, W \rangle = x_1 x_2 + y_1 y_2.$$

Similarly, we may define a Hermitian inner product $(\cdot, \cdot)$ in $\mathbb{C}$ by

$$(z, w) = z\overline{w}.$$

The term Hermitian is used to describe the fact that $(\cdot,\cdot)$ is not symmetric, but rather satisfies the relation

$$(z,w) = \overline{(w,z)} \qquad \text{for all } z,w \in \mathbb{C}.$$

Show that

$$\langle z,w\rangle = \frac{1}{2}[(z,w)+(w,z)] = \text{Re}(z,w),$$

where we use the usual identification $z = x+iy \in \mathbb{C}$ with $(x,y) \in \mathbb{R}^2$.

3. With $\omega = se^{i\varphi}$, where $s \geq 0$ and $\varphi \in \mathbb{R}$, solve the equation $z^n = \omega$ in $\mathbb{C}$ where n is a natural number. How many solutions are there?

4. Show that it is impossible to define a total ordering on $\mathbb{C}$. In other words, one cannot find a relation $\succ$ between complex numbers so that:

(i) For any two complex numbers z, w, one and only one of the following is true: $z \succ w$, $w \succ z$ or $z = w$.

(ii) For all $z_1, z_2, z_3 \in \mathbb{C}$ the relation $z_1 \succ z_2$ implies $z_1 + z_3 \succ z_2 + z_3$.

(iii) Moreover, for all $z_1, z_2, z_3 \in \mathbb{C}$ with $z_3 \succ 0$, then $z_1 \succ z_2$ implies $z_1 z_3 \succ z_2 z_3$.

[Hint: First check if $i \succ 0$ is possible.]

5. A set Ω is said to be **pathwise connected** if any two points in Ω can be joined by a (piecewise-smooth) curve entirely contained in Ω. The purpose of this exercise is to prove that an *open* set Ω is pathwise connected if and only if Ω is connected.

(a) Suppose first that Ω is open and pathwise connected, and that it can be written as $\Omega = \Omega_1 \cup \Omega_2$ where Ω_1 and Ω_2 are disjoint non-empty open sets. Choose two points $w_1 \in \Omega_1$ and $w_2 \in \Omega_2$ and let γ denote a curve in Ω joining w_1 to w_2. Consider a parametrization $z : [0,1] \to \Omega$ of this curve with $z(0) = w_1$ and $z(1) = w_2$, and let

$$t^* = \sup_{0\leq t\leq 1} \{t : z(s) \in \Omega_1 \quad \text{for all } 0 \leq s < t\}.$$

Arrive at a contradiction by considering the point $z(t^*)$.

(b) Conversely, suppose that Ω is open and connected. Fix a point $w \in \Omega$ and let $\Omega_1 \subset \Omega$ denote the set of all points that can be joined to w by a curve contained in Ω. Also, let $\Omega_2 \subset \Omega$ denote the set of all points that cannot be joined to w by a curve in Ω. Prove that both Ω_1 and Ω_2 are open, disjoint and their union is Ω. Finally, since Ω_1 is non-empty (why?) conclude that $\Omega = \Omega_1$ as desired.

The proof actually shows that the regularity and type of curves we used to define pathwise connectedness can be relaxed without changing the equivalence between the two definitions when Ω is open. For instance, we may take all curves to be continuous, or simply polygonal lines.[2]

6. Let Ω be an open set in $\mathbb{C}$ and $z \in \Omega$. The **connected component** (or simply the **component**) of Ω containing z is the set $\mathcal{C}_z$ of all points w in Ω that can be joined to z by a curve entirely contained in Ω.

(a) Check first that $\mathcal{C}_z$ is open and connected. Then, show that $w \in \mathcal{C}_z$ defines an equivalence relation, that is: (i) $z \in \mathcal{C}_z$, (ii) $w \in \mathcal{C}_z$ implies $z \in \mathcal{C}_w$, and (iii) if $w \in \mathcal{C}_z$ and $z \in \mathcal{C}_\zeta$, then $w \in \mathcal{C}_\zeta$.

Thus Ω is the union of all its connected components, and two components are either disjoint or coincide.

(b) Show that Ω can have only countably many distinct connected components.

(c) Prove that if Ω is the complement of a compact set, then Ω has only one unbounded component.

[Hint: For (b), one would otherwise obtain an uncountable number of disjoint open balls. Now, each ball contains a point with rational coordinates. For (c), note that the complement of a large disc containing the compact set is connected.]

7. The family of mappings introduced here plays an important role in complex analysis. These mappings, sometimes called **Blaschke factors**, will reappear in various applications in later chapters.

(a) Let z, w be two complex numbers such that $\overline{z}w \neq 1$. Prove that

$$\left|\frac{w-z}{1-\overline{w}z}\right| < 1 \qquad \text{if } |z| < 1 \text{ and } |w| < 1,$$

and also that

$$\left|\frac{w-z}{1-\overline{w}z}\right| = 1 \qquad \text{if } |z| = 1 \text{ or } |w| = 1.$$

[Hint: Why can one assume that z is real? It then suffices to prove that

$$(r-w)(r-\overline{w}) \leq (1-rw)(1-r\overline{w})$$

with equality for appropriate r and $|w|$.]

(b) Prove that for a fixed w in the unit disc $\mathbb{D}$, the mapping

$$F : z \mapsto \frac{w-z}{1-\overline{w}z}$$

satisfies the following conditions:

[2] A polygonal line is a piecewise-smooth curve which consists of finitely many straight line segments.

(i) F maps the unit disc to itself (that is, $F : \mathbb{D} \to \mathbb{D}$), and is holomorphic.

(ii) F interchanges 0 and w, namely $F(0) = w$ and $F(w) = 0$.

(iii) $|F(z)| = 1$ if $|z| = 1$.

(iv) $F : \mathbb{D} \to \mathbb{D}$ is bijective. [Hint: Calculate $F \circ F$.]

8. Suppose U and V are open sets in the complex plane. Prove that if $f : U \to V$ and $g : V \to \mathbb{C}$ are two functions that are differentiable (in the real sense, that is, as functions of the two real variables x and y), and $h = g \circ f$, then

$$\frac{\partial h}{\partial z} = \frac{\partial g}{\partial z}\frac{\partial f}{\partial z} + \frac{\partial g}{\partial \overline{z}}\frac{\partial \overline{f}}{\partial z}$$

and

$$\frac{\partial h}{\partial \overline{z}} = \frac{\partial g}{\partial z}\frac{\partial f}{\partial \overline{z}} + \frac{\partial g}{\partial \overline{z}}\frac{\partial \overline{f}}{\partial \overline{z}}.$$

This is the complex version of the chain rule.

9. Show that in polar coordinates, the Cauchy-Riemann equations take the form

$$\frac{\partial u}{\partial r} = \frac{1}{r}\frac{\partial v}{\partial \theta} \quad \text{and} \quad \frac{1}{r}\frac{\partial u}{\partial \theta} = -\frac{\partial v}{\partial r}.$$

Use these equations to show that the logarithm function defined by

$$\log z = \log r + i\theta \qquad \text{where } z = re^{i\theta} \text{ with } -\pi < \theta < \pi$$

is holomorphic in the region $r > 0$ and $-\pi < \theta < \pi$.

10. Show that

$$4\frac{\partial}{\partial z}\frac{\partial}{\partial \overline{z}} = 4\frac{\partial}{\partial \overline{z}}\frac{\partial}{\partial z} = \triangle,$$

where $\triangle$ is the **Laplacian**

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

11. Use Exercise 10 to prove that if f is holomorphic in the open set Ω, then the real and imaginary parts of f are **harmonic**; that is, their Laplacian is zero.

12. Consider the function defined by

$$f(x + iy) = \sqrt{|x||y|}, \qquad \text{whenever } x, y \in \mathbb{R}.$$

Show that f satisfies the Cauchy-Riemann equations at the origin, yet f is not holomorphic at 0.

13. Suppose that f is holomorphic in an open set Ω. Prove that in any one of the following cases:

(a) $\text{Re}(f)$ is constant;

(b) $\text{Im}(f)$ is constant;

(c) $|f|$ is constant;

one can conclude that f is constant.

14. Suppose $\{a_n\}_{n=1}^N$ and $\{b_n\}_{n=1}^N$ are two finite sequences of complex numbers. Let $B_k = \sum_{n=1}^k b_n$ denote the partial sums of the series $\sum b_n$ with the convention $B_0 = 0$. Prove the **summation by parts** formula

$$\sum_{n=M}^{N} a_n b_n = a_N B_N - a_M B_{M-1} - \sum_{n=M}^{N-1} (a_{n+1} - a_n) B_n.$$

15. Abel's theorem. Suppose $\sum_{n=1}^\infty a_n$ converges. Prove that

$$\lim_{r\to 1,\ r<1} \sum_{n=1}^{\infty} r^n a_n = \sum_{n=1}^{\infty} a_n.$$

[Hint: Sum by parts.] In other words, if a series converges, then it is Abel summable with the same limit. For the precise definition of these terms, and more information on summability methods, we refer the reader to Book I, Chapter 2.

16. Determine the radius of convergence of the series $\sum_{n=1}^\infty a_n z^n$ when:

(a) $a_n = (\log n)^2$

(b) $a_n = n!$

(c) $a_n = \frac{n^2}{4^n + 3n}$

(d) $a_n = (n!)^3/(3n)!$ [Hint: Use Stirling's formula, which says that $n! \sim c n^{n+\frac{1}{2}} e^{-n}$ for some $c > 0$..]

(e) Find the radius of convergence of the **hypergeometric series**

$$F(\alpha, \beta, \gamma; z) = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha+1)\cdots(\alpha+n-1)\beta(\beta+1)\cdots(\beta+n-1)}{n!\gamma(\gamma+1)\cdots(\gamma+n-1)} z^n.$$

Here $\alpha, \beta \in \mathbb{C}$ and $\gamma \neq 0, -1, -2, \ldots$.

(f) Find the radius of convergence of the Bessel function of order r:

$$J_r(z) = \left(\frac{z}{2}\right)^r \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+r)!} \left(\frac{z}{2}\right)^{2n},$$

where r is a positive integer.

17. Show that if $\{a_n\}_{n=0}^{\infty}$ is a sequence of non-zero complex numbers such that

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = L,$$

then

$$\lim_{n\to\infty} |a_n|^{1/n} = L.$$

In particular, this exercise shows that when applicable, the ratio test can be used to calculate the radius of convergence of a power series.

18. Let f be a power series centered at the origin. Prove that f has a power series expansion around any point in its disc of convergence.
[Hint: Write $z = z_0 + (z - z_0)$ and use the binomial expansion for z^n.]

19. Prove the following:

(a) The power series $\sum nz^n$ does not converge on any point of the unit circle.

(b) The power series $\sum z^n/n^2$ converges at every point of the unit circle.

(c) The power series $\sum z^n/n$ converges at every point of the unit circle except $z = 1$. [Hint: Sum by parts.]

20. Expand $(1-z)^{-m}$ in powers of z. Here m is a fixed positive integer. Also, show that if

$$(1-z)^{-m} = \sum_{n=0}^{\infty} a_n z^n,$$

then one obtains the following asymptotic relation for the coefficients:

$$a_n \sim \frac{1}{(m-1)!}\, n^{m-1} \quad \text{as } n \to \infty.$$

21. Show that for $|z| < 1$, one has

$$\frac{z}{1-z^2} + \frac{z^2}{1-z^4} + \cdots + \frac{z^{2^n}}{1-z^{2^{n+1}}} + \cdots = \frac{z}{1-z},$$

and

$$\frac{z}{1+z} + \frac{2z^2}{1+z^2} + \cdots + \frac{2^k z^{2^k}}{1+z^{2^k}} + \cdots = \frac{z}{1-z}.$$

Justify any change in the order of summation.

[Hint: Use the dyadic expansion of an integer and the fact that $2^{k+1} - 1 = 1 + 2 + 2^2 + \cdots + 2^k$.]

22. Let $\mathbb{N} = \{1, 2, 3, \ldots\}$ denote the set of positive integers. A subset $S \subset \mathbb{N}$ is said to be in arithmetic progression if

$$S = \{a, a+d, a+2d, a+3d, \ldots\}$$

where $a, d \in \mathbb{N}$. Here d is called the step of S.

Show that $\mathbb{N}$ cannot be partitioned into a finite number of subsets that are in arithmetic progression with distinct steps (except for the trivial case $a = d = 1$).

[Hint: Write $\sum_{n\in\mathbb{N}} z^n$ as a sum of terms of the type $\frac{z^a}{1-z^d}$.]

23. Consider the function f defined on $\mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0\,, \\ e^{-1/x^2} & \text{if } x > 0. \end{cases}$$

Prove that f is indefinitely differentiable on $\mathbb{R}$, and that $f^{(n)}(0) = 0$ for all $n \geq 1$. Conclude that f does not have a converging power series expansion $\sum_{n=0}^{\infty} a_n x^n$ for x near the origin.

24. Let γ be a smooth curve in $\mathbb{C}$ parametrized by $z(t) : [a, b] \to \mathbb{C}$. Let γ^- denote the curve with the same image as γ but with the reverse orientation. Prove that for any continuous function f on γ

$$\int_\gamma f(z)\, dz = -\int_{\gamma^-} f(z)\, dz.$$

25. The next three calculations provide some insight into Cauchy's theorem, which we treat in the next chapter.

(a) Evaluate the integrals

$$\int_\gamma z^n\, dz$$

for all integers n. Here γ is any circle centered at the origin with the positive (counterclockwise) orientation.

(b) Same question as before, but with γ any circle not containing the origin.

(c) Show that if $|a| < r < |b|$, then

$$\int_\gamma \frac{1}{(z-a)(z-b)}\, dz = \frac{2\pi i}{a-b},$$

where γ denotes the circle centered at the origin, of radius r, with the positive orientation.

26. Suppose f is continuous in a region Ω. Prove that any two primitives of f (if they exist) differ by a constant.

2 Cauchy's Theorem and Its Applications

> The solution of a large number of problems can be reduced, in the last analysis, to the evaluation of definite integrals; thus mathematicians have been much occupied with this task... However, among many results obtained, a number were initially discovered by the aid of a type of induction based on the passage from real to imaginary. Often passage of this kind led directly to remarkable results. Nevertheless this part of the theory, as has been observed by Laplace, is subject to various difficulties...
>
> After having reflected on this subject and brought together various results mentioned above, I hope to establish the passage from the real to the imaginary based on a direct and rigorous analysis; my researches have thus led me to the method which is the object of this memoir...
>
> *A. L. Cauchy,* 1827

In the previous chapter, we discussed several preliminary ideas in complex analysis: open sets in $\mathbb{C}$, holomorphic functions, and integration along curves. The first remarkable result of the theory exhibits a deep connection between these notions. Loosely stated, Cauchy's theorem says that if f is holomorphic in an open set Ω and $\gamma \subset \Omega$ is a closed curve whose interior is also contained in Ω then

$$\int_\gamma f(z)\,dz = 0. \tag{1}$$

Many results that follow, and in particular the calculus of residues, are related in one way or another to this fact.

A precise and general formulation of Cauchy's theorem requires defining unambiguously the "interior" of a curve, and this is not always an easy task. At this early stage of our study, we shall make use of the device of limiting ourselves to regions whose boundaries are curves that are "toy contours." As the name suggests, these are closed curves whose visualization is so simple that the notion of their interior will be unam-

biguous, and the proof of Cauchy's theorem in this setting will be quite direct. For many applications, it will suffice to restrict ourselves to these types of curves. At a later stage, we take up the questions related to more general curves, their interiors, and the extended form of Cauchy's theorem.

Our initial version of Cauchy's theorem begins with the observation that it suffices that f have a primitive in Ω, by Corollary 3.3 in Chapter 1. The existence of such a primitive for toy contours will follow from a theorem of Goursat (which is itself a simple special case)[1] that asserts that if f is holomorphic in an open set that contains a triangle T and its interior, then

$$\int_T f(z)\, dz = 0.$$

It is noteworthy that this simple case of Cauchy's theorem suffices to prove some of its more complicated versions. From there, we can prove the existence of primitives in the interior of some simple regions, and therefore prove Cauchy's theorem in that setting. As a first application of this viewpoint, we evaluate several real integrals by using appropriate toy contours.

The above ideas also lead us to a central result of this chapter, the Cauchy integral formula; this states that if f is holomorphic in an open set containing a circle C and its interior, then for all z inside C,

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta.$$

Differentiation of this identity yields other integral formulas, and in particular we obtain the regularity of holomorphic functions. This is remarkable, since holomorphicity assumed only the existence of the first derivative, and yet we obtain as a consequence the existence of derivatives of all orders. (An analogous statement is decisively false in the case of real variables!)

The theory developed up to that point already has a number of noteworthy consequences:

- The property at the base of "analytic continuation," namely that a holomorphic function is determined by its restriction to any open subset of its domain of definition. This is a consequence of the fact that holomorphic functions have power series expansions.

[1]Goursat's result came after Cauchy's theorem, and its interest is the technical fact that its proof requires only the existence of the complex derivative at each point, and not its continuity. For the earlier proof, see Exercise 5.

- Liouville's theorem, which yields a quick proof of the fundamental theorem of algebra.
- Morera's theorem, which gives a simple integral characterization of holomorphic functions, and shows that these functions are preserved under uniform limits.

1 Goursat's theorem

Corollary 3.3 in the previous chapter says that if f has a primitive in an open set Ω, then

$$\int_\gamma f(z)\,dz = 0$$

for any closed curve γ in Ω. Conversely, if we can show that the above relation holds for some types of curves γ, then a primitive will exist. Our starting point is Goursat's theorem, from which in effect we shall deduce most of the other results in this chapter.

Theorem 1.1 *If Ω is an open set in $\mathbb{C}$, and $T \subset \Omega$ a triangle whose interior is also contained in Ω, then*

$$\int_T f(z)\,dz = 0$$

whenever f is holomorphic in Ω.

Proof. We call $T^{(0)}$ our original triangle (with a fixed orientation which we choose to be positive), and let $d^{(0)}$ and $p^{(0)}$ denote the diameter and perimeter of $T^{(0)}$, respectively. The first step in our construction consists of bisecting each side of the triangle and connecting the midpoints. This creates four new smaller triangles, denoted $T_1^{(1)}, T_2^{(1)}, T_3^{(1)}$, and $T_4^{(1)}$ that are similar to the original triangle. The construction and orientation of each triangle are illustrated in Figure 1. The orientation is chosen to be consistent with that of the original triangle, and so after cancellations arising from integrating over the same side in two opposite directions, we have

$$\int_{T^{(0)}} f(z)\,dz = \int_{T_1^{(1)}} f(z)\,dz + \int_{T_2^{(1)}} f(z)\,dz + \int_{T_3^{(1)}} f(z)\,dz + \int_{T_4^{(1)}} f(z)\,dz. \tag{2}$$

For some j we must have

$$\left|\int_{T^{(0)}} f(z)\,dz\right| \leq 4\left|\int_{T_j^{(1)}} f(z)\,dz\right|,$$

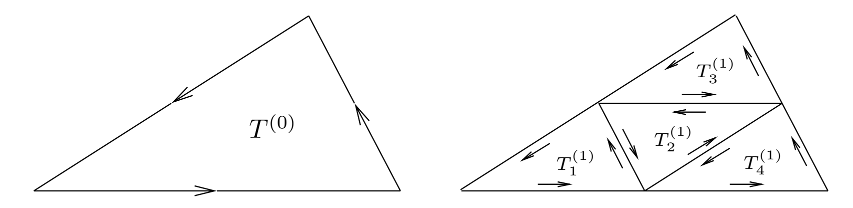


Figure 1. Bisection of $T^{(0)}$

for otherwise (2) would be contradicted. We choose a triangle that satisfies this inequality, and rename it $T^{(1)}$. Observe that if $d^{(1)}$ and $p^{(1)}$ denote the diameter and perimeter of $T^{(1)}$, respectively, then $d^{(1)} = (1/2)d^{(0)}$ and $p^{(1)} = (1/2)p^{(0)}$. We now repeat this process for the triangle $T^{(1)}$, bisecting it into four smaller triangles. Continuing this process, we obtain a sequence of triangles

$$T^{(0)}, T^{(1)}, \ldots, T^{(n)}, \ldots$$

with the properties that

$$\left|\int_{T^{(0)}} f(z)\, dz\right| \leq 4^n \left|\int_{T^{(n)}} f(z)\, dz\right|$$

and

$$d^{(n)} = 2^{-n} d^{(0)}, \qquad p^{(n)} = 2^{-n} p^{(0)}$$

where $d^{(n)}$ and $p^{(n)}$ denote the diameter and perimeter of $T^{(n)}$, respectively. We also denote by $\mathcal{T}^{(n)}$ the *solid* closed triangle with boundary $T^{(n)}$, and observe that our construction yields a sequence of nested compact sets

$$\mathcal{T}^{(0)} \supset \mathcal{T}^{(1)} \supset \cdots \supset \mathcal{T}^{(n)} \supset \cdots$$

whose diameter goes to 0. By Proposition 1.4 in Chapter 1, there exists a unique point z_0 that belongs to all the solid triangles $\mathcal{T}^{(n)}$. Since f is holomorphic at z_0 we can write

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \psi(z)(z - z_0)\,,$$

where $\psi(z) \to 0$ as $z \to z_0$. Since the constant $f(z_0)$ and the linear function $f'(z_0)(z - z_0)$ have primitives, we can integrate the above equality using Corollary 3.3 in the previous chapter, and obtain

$$\int_{T^{(n)}} f(z)\, dz = \int_{T^{(n)}} \psi(z)(z - z_0)\, dz. \tag{3}$$

Now z_0 belongs to the closure of the solid triangle $\mathcal{T}^{(n)}$ and z to its boundary, so we must have $|z - z_0| \leq d^{(n)}$, and using (3) we get, by (iii) in Proposition 3.1 of the previous chapter, the estimate

$$\left| \int_{T^{(n)}} f(z)\, dz \right| \leq \epsilon_n d^{(n)} p^{(n)},$$

where $\epsilon_n = \sup_{z \in T^{(n)}} |\psi(z)| \to 0$ as $n \to \infty$. Therefore

$$\left| \int_{T^{(n)}} f(z)\, dz \right| \leq \epsilon_n 4^{-n} d^{(0)} p^{(0)} \,,$$

which yields our final estimate

$$\left| \int_{T^{(0)}} f(z)\, dz \right| \leq 4^n \left| \int_{T^{(n)}} f(z)\, dz \right| \leq \epsilon_n d^{(0)} p^{(0)}.$$

Letting $n \to \infty$ concludes the proof since $\epsilon_n \to 0$.

Corollary 1.2 *If f is holomorphic in an open set Ω that contains a rectangle R and its interior, then*

$$\int_R f(z)\, dz = 0.$$

This is immediate since we first choose an orientation as in Figure 2 and note that

$$\int_R f(z)\, dz = \int_{T_1} f(z)\, dz + \int_{T_2} f(z)\, dz.$$

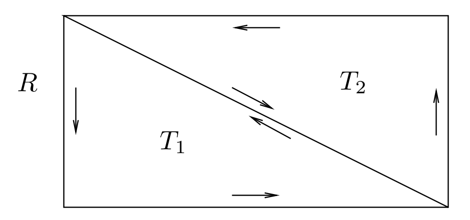


Figure 2. A rectangle as the union of two triangles

2 Local existence of primitives and Cauchy's theorem in a disc

We first prove the existence of primitives in a disc as a consequence of Goursat's theorem.

Theorem 2.1 *A holomorphic function in an open disc has a primitive in that disc.*

Proof. After a translation, we may assume without loss of generality that the disc, say D, is centered at the origin. Given a point $z \in D$, consider the piecewise-smooth curve that joins 0 to z first by moving in the horizontal direction from 0 to $\tilde{z}$ where $\tilde{z} = \operatorname{Re}(z)$, and then in the vertical direction from $\tilde{z}$ to z. We choose the orientation from 0 to z, and denote this polygonal line (which consists of at most two segments) by γ_z, as shown on Figure 3.

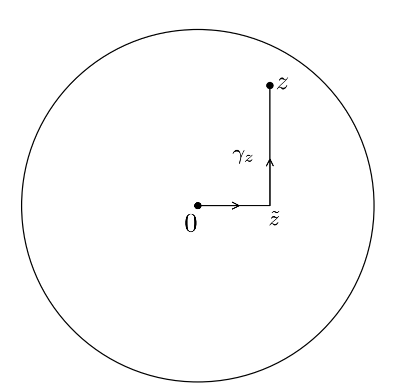


Figure 3. The polygonal line γ_z

Define

$$F(z) = \int_{\gamma_z} f(w)\, dw.$$

The choice of γ_z gives an unambiguous definition of the function $F(z)$. We contend that F is holomorphic in D and $F'(z) = f(z)$. To prove this, fix $z \in D$ and let $h \in \mathbb{C}$ be so small that $z + h$ also belongs to the disc. Now consider the difference

$$F(z+h) - F(z) = \int_{\gamma_{z+h}} f(w)\, dw - \int_{\gamma_z} f(w)\, dw.$$

The function f is first integrated along γ_{z+h} with the original orientation, and then along γ_z with the reverse orientation (because of the minus sign in front of the second integral). This corresponds to (a) in Figure 4. Since we integrate f over the line segment starting at the origin in two opposite directions, it cancels, leaving us with the contour in (b). Then, we complete the square and triangle as shown in (c), so that after an application of Goursat's theorem for triangles and rectangles we are left with the line segment from z to $z+h$ as given in (d).

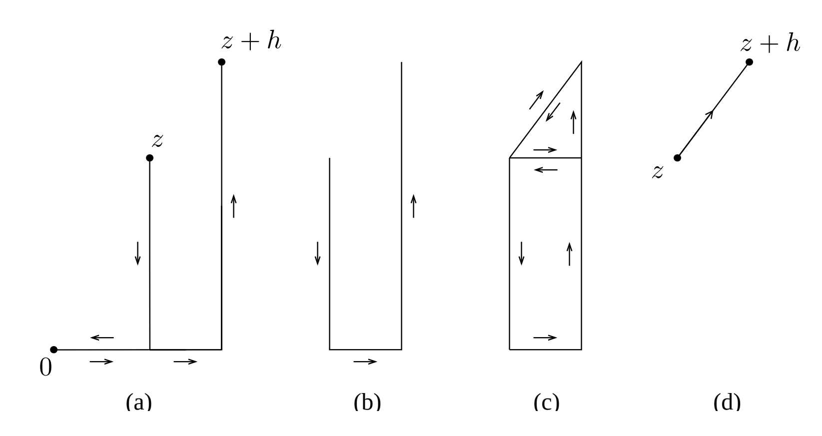


Figure 4. Relation between the polygonal lines γ_z and γ_{z+h}

Hence the above cancellations yield

$$F(z+h) - F(z) = \int_\eta f(w)\, dw$$

where η is the straight line segment from z to $z+h$. Since f is continuous at z we can write

$$f(w) = f(z) + \psi(w)$$

where $\psi(w) \to 0$ as $w \to z$. Therefore

(4)

$$F(z+h) - F(z) = \int_\eta f(z)\, dw + \int_\eta \psi(w)\, dw = f(z) \int_\eta dw + \int_\eta \psi(w)\, dw.$$

On the one hand, the constant 1 has w as a primitive, so the first integral is simply h by an application of Theorem 3.2 in Chapter 1. On the other

hand, we have the following estimate:

$$\left|\int_\eta \psi(w)\, dw\right| \leq \sup_{w\in\eta} |\psi(w)|\, |h|.$$

Since the supremum above goes to 0 as h tends to 0, we conclude from equation (4) that

$$\lim_{h\to 0} \frac{F(z+h)-F(z)}{h} = f(z)\,,$$

thereby proving that F is a primitive for f on the disc.

This theorem says that locally, every holomorphic function has a primitive. It is crucial to realize, however, that the theorem is true not only for arbitrary discs, but also for other sets as well. We shall return to this point shortly in our discussion of "toy contours."

Theorem 2.2 (Cauchy's theorem for a disc) *If f is holomorphic in a disc, then*

$$\int_\gamma f(z)\, dz = 0$$

for any closed curve γ in that disc.

Proof. Since f has a primitive, we can apply Corollary 3.3 of Chapter 1.

Corollary 2.3 *Suppose f is holomorphic in an open set containing the circle C and its interior. Then*

$$\int_C f(z)\, dz = 0.$$

Proof. Let D be the disc with boundary circle C. Then there exists a slightly larger disc D' which contains D and so that f is holomorphic on D'. We may now apply Cauchy's theorem in D' to conclude that $\int_C f(z)\, dz = 0$.

In fact, the proofs of the theorem and its corollary apply whenever we can define without ambiguity the "interior" of a contour, and construct appropriate polygonal paths in an open neighborhood of that contour and its interior. In the case of the circle, whose interior is the disc, there was no problem since the geometry of the disc made it simple to travel horizontally and vertically inside it.

The following definition is loosely stated, although its applications will be clear and unambiguous. We call a **toy contour** any closed curve where the notion of interior is obvious, and a construction similar to that in Theorem 2.1 is possible in a neighborhood of the curve and its interior. Its positive orientation is that for which the interior is to the left as we travel along the toy contour. This is consistent with the definition of the positive orientation of a circle. For example, circles, triangles, and rectangles are toy contours, since in each case we can modify (and actually copy) the argument given previously.

Another important example of a toy contour is the "keyhole" Γ (illustrated in Figure 5), which we shall put to use in the proof of the Cauchy integral formula. It consists of two almost complete circles, one large

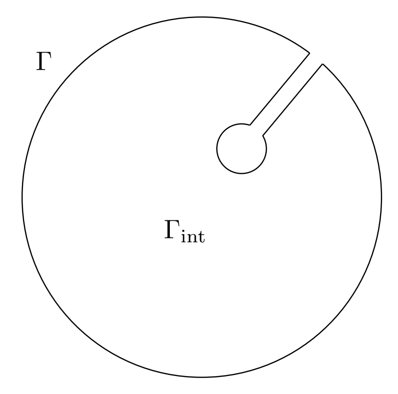


Figure 5. The keyhole contour

and one small, connected by a narrow corridor. The interior of Γ, which we denote by Γ_{int}, is clearly that region enclosed by the curve, and can be given precise meaning with enough work. We fix a point z_0 in that interior. If f is holomorphic in a neighborhood of Γ and its interior, then it is holomorphic in the inside of a slightly larger keyhole, say Λ, whose interior Λ_{int} contains $\Gamma \cup \Gamma_{\text{int}}$. If $z \in \Lambda_{\text{int}}$, let γ_z denote any curve contained inside Λ_{int} connecting z_0 to z, and which consists of finitely many horizontal or vertical segments (as in Figure 6). If η_z is any other such curve, the rectangle version of Goursat's theorem (Corollary 1.2) implies that

$$\int_{\gamma_z} f(w)\,dw = \int_{\eta_z} f(w)\,dw\,,$$

and we may therefore define F unambiguously in Λ_{int}.

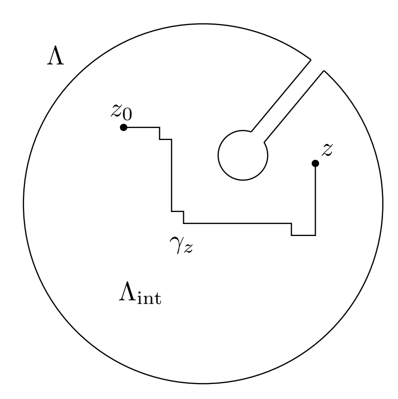


Figure 6. A curve γ_z

Arguing as above allows us to show that F is a primitive of f in Λ_{int} and therefore $\int_\Gamma f(z)\,dz = 0$.

The important point is that for a toy contour γ we easily have that

$$\int_\gamma f(z)\,dz = 0\,,$$

whenever f is holomorphic in an open set that contains the contour γ and its interior.

Other examples of toy contours which we shall encounter in applications and for which Cauchy's theorem and its corollary also hold are given in Figure 7.

While Cauchy's theorem for toy contours is sufficient for most applications we deal with, the question still remains as to what happens for more general curves. We take up this matter in Appendix B, where we prove Jordan's theorem for piecewise-smooth curves. This theorem states that a simple closed piecewise-smooth curve has a well defined interior that is "simply connected." As a consequence, we find that even in this more general situation, Cauchy's theorem holds.

3 Evaluation of some integrals

Here we take up the idea that originally motivated Cauchy. We shall show by several examples how some integrals may be evaluated by the use of his theorem. A more systematic approach, in terms of the calculus of residues, may be found in the next chapter.

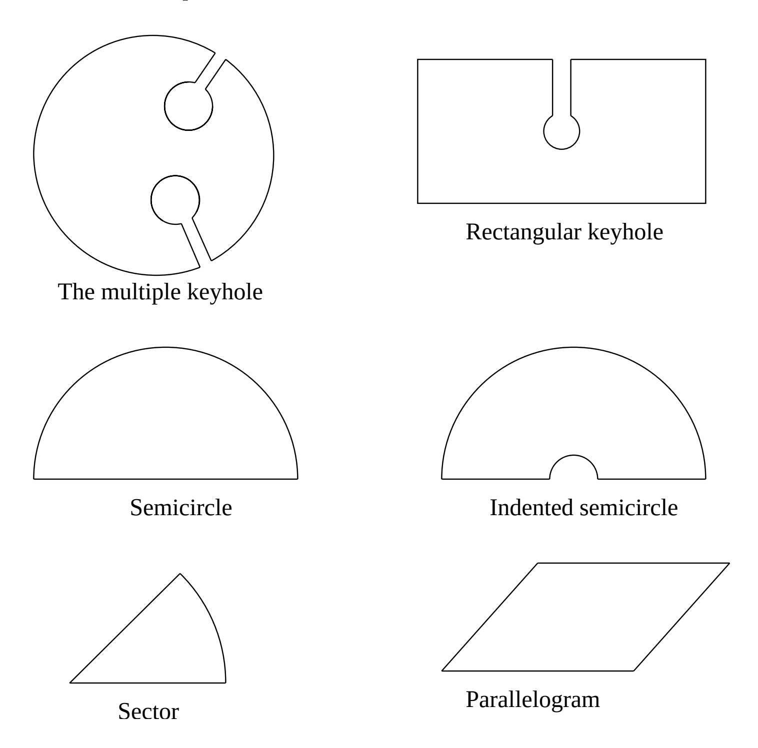


Figure 7. Examples of toy contours

EXAMPLE 1. We show that if $\xi \in \mathbb{R}$, then

$$e^{-\pi\xi^2} = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx. \tag{5}$$

This gives a new proof of the fact that $e^{-\pi x^2}$ is its own Fourier transform, a fact we proved in Theorem 1.4 of Chapter 5 in Book I.

If $\xi = 0$, the formula is precisely the known integral[2]

$$1 = \int_{-\infty}^{\infty} e^{-\pi x^2}\, dx.$$

Now suppose that $\xi > 0$, and consider the function $f(z) = e^{-\pi z^2}$, which is entire, and in particular holomorphic in the interior of the toy contour γ_R depicted in Figure 8.

[2]An alternate derivation follows from the fact that $\Gamma(1/2) = \sqrt{\pi}$, where Γ is the gamma function in Chapter 6.

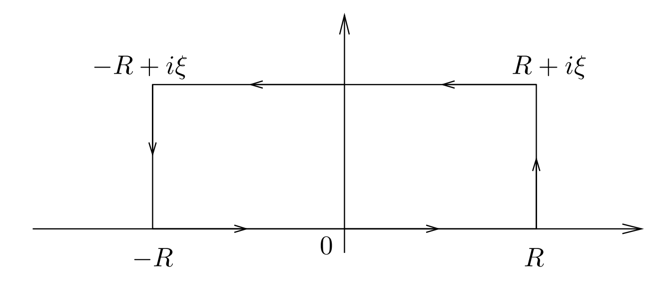


Figure 8. The contour γ_R in Example 1

The contour γ_R consists of a rectangle with vertices $R, R+i\xi, -R+i\xi, -R$ and the positive counterclockwise orientation. By Cauchy's theorem,

$$\int_{\gamma_R} f(z)\, dz = 0. \tag{6}$$

The integral over the real segment is simply

$$\int_{-R}^{R} e^{-\pi x^2}\, dx\,,$$

which converges to 1 as $R \to \infty$. The integral on the vertical side on the right is

$$I(R) = \int_0^{\xi} f(R+iy) i\, dy = \int_0^{\xi} e^{-\pi(R^2+2iRy-y^2)} i\, dy.$$

This integral goes to 0 as $R \to \infty$ since ξ is fixed and we may estimate it by

$$|I(R)| \leq Ce^{-\pi R^2}.$$

Similarly, the integral over the vertical segment on the left also goes to 0 as $R \to \infty$ for the same reasons. Finally, the integral over the horizontal segment on top is

$$\int_{R}^{-R} e^{-\pi(x+i\xi)^2}\, dx = -e^{\pi\xi^2} \int_{-R}^{R} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx.$$

Therefore, we find in the limit as $R \to \infty$ that (6) gives

$$0 = 1 - e^{\pi\xi^2} \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx,$$

and our desired formula is established. In the case $\xi < 0$, we then consider the symmetric rectangle, in the lower half-plane.

The technique of shifting the contour of integration, which was used in the previous example, has many other applications. Note that the original integral (5) is taken over the real line, which by an application of Cauchy's theorem is then shifted upwards or downwards (depending on the sign of ξ) in the complex plane.

EXAMPLE 2. Another classical example is

$$\int_0^\infty \frac{1-\cos x}{x^2}\,dx = \frac{\pi}{2}.$$

Here we consider the function $f(z) = (1-e^{iz})/z^2$, and we integrate over the indented semicircle in the upper half-plane positioned on the x-axis, as shown in Figure 9.

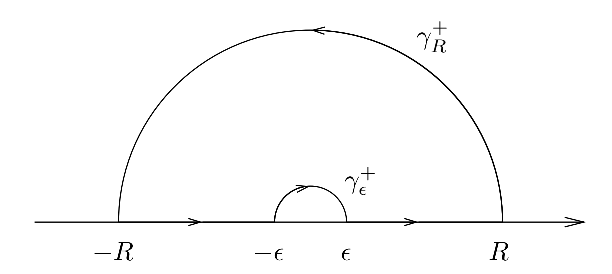


Figure 9. The indented semicircle of Example 2

If we denote by γ_ϵ^+ and γ_R^+ the semicircles of radii ϵ and R with negative and positive orientations respectively, Cauchy's theorem gives

$$\int_{-R}^{-\epsilon} \frac{1-e^{ix}}{x^2}\,dx + \int_{\gamma_\epsilon^+} \frac{1-e^{iz}}{z^2}\,dz + \int_\epsilon^R \frac{1-e^{ix}}{x^2}\,dx + \int_{\gamma_R^+} \frac{1-e^{iz}}{z^2}\,dz = 0.$$

First we let $R \to \infty$ and observe that

$$\left|\frac{1-e^{iz}}{z^2}\right| \le \frac{2}{|z|^2},$$

so the integral over γ_R^+ goes to zero. Therefore

$$\int_{|x|\ge\epsilon} \frac{1-e^{ix}}{x^2}\,dx = -\int_{\gamma_\epsilon^+} \frac{1-e^{iz}}{z^2}\,dz.$$

Next, note that

$$f(z) = \frac{-iz}{z^2} + E(z)$$

where $E(z)$ is bounded as $z \to 0$, while on γ_ϵ^+ we have $z = \epsilon e^{i\theta}$ and $dz = i\epsilon e^{i\theta} d\theta$. Thus

$$\int_{\gamma_\epsilon^+} \frac{1 - e^{iz}}{z^2}\, dz \to \int_\pi^0 (-ii)\, d\theta = -\pi \qquad \text{as } \epsilon \to 0.$$

Taking real parts then yields

$$\int_{-\infty}^{\infty} \frac{1 - \cos x}{x^2}\, dx = \pi.$$

Since the integrand is even, the desired formula is proved.

4 Cauchy's integral formulas

Representation formulas, and in particular integral representation formulas, play an important role in mathematics, since they allow us to recover a function on a large set from its behavior on a smaller set. For example, we saw in Book I that a solution of the steady-state heat equation in the disc was completely determined by its boundary values on the circle via a convolution with the Poisson kernel

$$u(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} P_r(\theta - \varphi) u(1, \varphi)\, d\varphi. \tag{7}$$

In the case of holomorphic functions, the situation is analogous, which is not surprising since the real and imaginary parts of a holomorphic function are harmonic.[3] Here, we will prove an integral representation formula in a manner that is independent of the theory of harmonic functions. In fact, it is also possible to recover the Poisson integral formula (7) as a consequence of the next theorem (see Exercises 11 and 12).

Theorem 4.1 *Suppose f is holomorphic in an open set that contains the closure of a disc D. If C denotes the boundary circle of this disc with the positive orientation, then*

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta \qquad \textit{for any point } z \in D.$$

[3]This fact is an immediate consequence of the Cauchy-Riemann equations. We refer the reader to Exercise 11 in Chapter 1.

Proof. Fix $z \in D$ and consider the "keyhole" $\Gamma_{\delta,\epsilon}$ which omits the point z as shown in Figure 10.

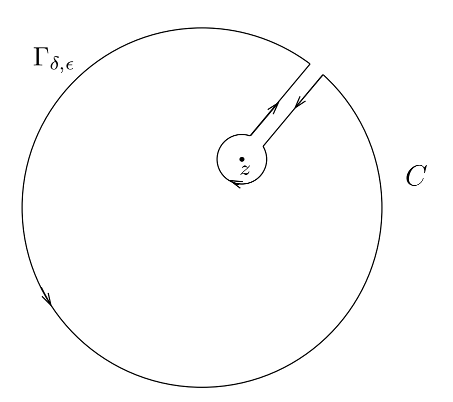


Figure 10. The keyhole $\Gamma_{\delta,\epsilon}$

Here δ is the width of the corridor, and ϵ the radius of the small circle centered at z. Since the function $F(\zeta) = f(\zeta)/(\zeta - z)$ is holomorphic away from the point $\zeta = z$, we have

$$\int_{\Gamma_{\delta,\epsilon}} F(\zeta)\, d\zeta = 0$$

by Cauchy's theorem for the chosen toy contour. Now we make the corridor narrower by letting δ tend to 0, and use the continuity of F to see that in the limit, the integrals over the two sides of the corridor cancel out. The remaining part consists of two curves, the large boundary circle C with the positive orientation, and a small circle C_ϵ centered at z of radius ϵ and oriented negatively, that is, clockwise. To see what happens to the integral over the small circle we write

$$F(\zeta) = \frac{f(\zeta) - f(z)}{\zeta - z} + \frac{f(z)}{\zeta - z} \tag{8}$$

and note that since f is holomorphic the first term on the right-hand side of (8) is bounded so that its integral over C_ϵ goes to 0 as $\epsilon \to 0$. To

conclude the proof, it suffices to observe that

$$\begin{aligned}\int_{C_\epsilon} \frac{f(z)}{\zeta - z}\, d\zeta &= f(z) \int_{C_\epsilon} \frac{d\zeta}{\zeta - z} \\ &= -f(z) \int_0^{2\pi} \frac{\epsilon i e^{-it}}{\epsilon e^{-it}}\, dt \\ &= -f(z) 2\pi i\,,\end{aligned}$$

so that in the limit we find

$$0 = \int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta - 2\pi i f(z)\,,$$

as was to be shown.

Remarks. Our earlier discussion of toy contours provides simple extensions of the Cauchy integral formula; for instance, if f is holomorphic in an open set that contains a (positively oriented) rectangle R and its interior, then

$$f(z) = \frac{1}{2\pi i} \int_R \frac{f(\zeta)}{\zeta - z}\, d\zeta\,,$$

whenever z belongs to the interior of R. To establish this result, it suffices to repeat the proof of Theorem 4.1 replacing the "circular" keyhole by a "rectangular" keyhole.

It should also be noted that the above integral vanishes when z is outside R, since in this case $F(\zeta) = f(\zeta)/(\zeta - z)$ is holomorphic inside R. Of course, a similar result also holds for the circle or any other toy contour.

As a corollary to the Cauchy integral formula, we arrive at a second remarkable fact about holomorphic functions, namely their regularity. We also obtain further integral formulas expressing the derivatives of f inside the disc in terms of the values of f on the boundary.

Corollary 4.2 *If f is holomorphic in an open set Ω, then f has infinitely many complex derivatives in Ω. Moreover, if $C \subset \Omega$ is a circle whose interior is also contained in Ω, then*

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta$$

for all z in the interior of C.

We recall that, as in the above theorem, we take the circle C to have positive orientation.

Proof. The proof is by induction on n, the case $n = 0$ being simply the Cauchy integral formula. Suppose that f has up to $n-1$ complex derivatives and that

$$f^{(n-1)}(z) = \frac{(n-1)!}{2\pi i}\int_C \frac{f(\zeta)}{(\zeta - z)^n}\,d\zeta.$$

Now for h small, the difference quotient for $f^{(n-1)}$ takes the form

$$\frac{f^{(n-1)}(z+h) - f^{(n-1)}(z)}{h} = \frac{(n-1)!}{2\pi i}\int_C f(\zeta)\frac{1}{h}\left[\frac{1}{(\zeta - z - h)^n} - \frac{1}{(\zeta - z)^n}\right] d\zeta. \tag{9}$$

We now recall that

$$A^n - B^n = (A - B)[A^{n-1} + A^{n-2}B + \cdots + AB^{n-2} + B^{n-1}].$$

With $A = 1/(\zeta - z - h)$ and $B = 1/(\zeta - z)$, we see that the term in brackets in equation (9) is equal to

$$\frac{h}{(\zeta - z - h)(\zeta - z)}[A^{n-1} + A^{n-2}B + \cdots + AB^{n-2} + B^{n-1}].$$

But observe that if h is small, then $z + h$ and z stay at a finite distance from the boundary circle C, so in the limit as h tends to 0, we find that the quotient converges to

$$\frac{(n-1)!}{2\pi i}\int_C f(\zeta)\left[\frac{1}{(\zeta - z)^2}\right]\left[\frac{n}{(\zeta - z)^{n-1}}\right] d\zeta = \frac{n!}{2\pi i}\int_C \frac{f(\zeta)}{(\zeta - z)^{n+1}}\,d\zeta\,,$$

which completes the induction argument and proves the theorem.

From now on, we call the formulas of Theorem 4.1 and Corollary 4.2 the **Cauchy integral formulas**.

Corollary 4.3 (Cauchy inequalities) *If f is holomorphic in an open set that contains the closure of a disc D centered at z_0 and of radius R, then*

$$|f^{(n)}(z_0)| \leq \frac{n!\|f\|_C}{R^n}\,,$$

where $\|f\|_C = \sup_{z\in C}|f(z)|$ denotes the supremum of $|f|$ on the boundary circle C.

Proof. Applying the Cauchy integral formula for $f^{(n)}(z_0)$, we obtain

$$\begin{aligned} |f^{(n)}(z_0)| &= \left| \frac{n!}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} \, d\zeta \right| \\ &= \frac{n!}{2\pi} \left| \int_0^{2\pi} \frac{f(z_0 + Re^{i\theta})}{(Re^{i\theta})^{n+1}} Rie^{i\theta} \, d\theta \right| \\ &\leq \frac{n!}{2\pi} \frac{\|f\|_C}{R^n} 2\pi. \end{aligned}$$

Another striking consequence of the Cauchy integral formula is its connection with power series. In Chapter 1, we proved that a power series is holomorphic in the interior of its disc of convergence, and promised a proof of a converse, which is the content of the next theorem.

Theorem 4.4 *Suppose f is holomorphic in an open set Ω. If D is a disc centered at z_0 and whose closure is contained in Ω, then f has a power series expansion at z_0*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

for all $z \in D$, and the coefficients are given by

$$a_n = \frac{f^{(n)}(z_0)}{n!} \qquad \text{for all } n \geq 0.$$

Proof. Fix $z \in D$. By the Cauchy integral formula, we have

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z} \, d\zeta \,, \tag{10}$$

where C denotes the boundary of the disc and $z \in D$. The idea is to write

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - z_0 - (z - z_0)} = \frac{1}{\zeta - z_0} \frac{1}{1 - \left(\frac{z - z_0}{\zeta - z_0}\right)}, \tag{11}$$

and use the geometric series expansion. Since $\zeta \in C$ and $z \in D$ is fixed, there exists $0 < r < 1$ such that

$$\left| \frac{z - z_0}{\zeta - z_0} \right| < r,$$

therefore

$$\frac{1}{1-\left(\frac{z-z_0}{\zeta-z_0}\right)}=\sum_{n=0}^{\infty}\left(\frac{z-z_0}{\zeta-z_0}\right)^n, \tag{12}$$

where the series converges uniformly for $\zeta \in C$. This allows us to interchange the infinite sum with the integral when we combine (10), (11), and (12), thereby obtaining

$$f(z)=\sum_{n=0}^{\infty}\left(\frac{1}{2\pi i}\int_C \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\,d\zeta\right)\cdot(z-z_0)^n.$$

This proves the power series expansion; further the use of the Cauchy integral formulas for the derivatives (or simply differentiation of the series) proves the formula for a_n.

Observe that since power series define indefinitely (complex) differentiable functions, the theorem gives another proof that a holomorphic function is automatically indefinitely differentiable.

Another important observation is that the power series expansion of f centered at z_0 converges in any disc, no matter how large, as long as its closure is contained in Ω. In particular, if f is entire (that is, holomorphic on all of $\mathbb{C}$), the theorem implies that f has a power series expansion around 0, say $f(z)=\sum_{n=0}^{\infty} a_n z^n$, that converges in all of $\mathbb{C}$.

Corollary 4.5 (Liouville's theorem) *If f is entire and bounded, then f is constant.*

Proof. It suffices to prove that $f'=0$, since $\mathbb{C}$ is connected, and we may then apply Corollary 3.4 in Chapter 1.

For each $z_0 \in \mathbb{C}$ and all $R>0$, the Cauchy inequalities yield

$$|f'(z_0)| \leq \frac{B}{R}$$

where B is a bound for f. Letting $R \to \infty$ gives the desired result.

As an application of our work so far, we can give an elegant proof of the fundamental theorem of algebra.

Corollary 4.6 *Every non-constant polynomial $P(z)=a_n z^n+\cdots+a_0$ with complex coefficients has a root in $\mathbb{C}$.*

Proof. If P has no roots, then $1/P(z)$ is a bounded holomorphic function. To see this, we can of course assume that $a_n \neq 0$, and write

$$\frac{P(z)}{z^n}=a_n+\left(\frac{a_{n-1}}{z}+\cdots+\frac{a_0}{z^n}\right)$$

whenever $z \neq 0$. Since each term in the parentheses goes to 0 as $|z| \to \infty$ we conclude that there exists $R > 0$ so that if $c = |a_n|/2$, then

$$|P(z)| \geq c|z|^n \quad \text{whenever } |z| > R.$$

In particular, P is bounded from below when $|z| > R$. Since P is continuous and has no roots in the disc $|z| \leq R$, it is bounded from below in that disc as well, thereby proving our claim.

By Liouville's theorem we then conclude that $1/P$ is constant. This contradicts our assumption that P is non-constant and proves the corollary.

Corollary 4.7 *Every polynomial $P(z) = a_n z^n + \cdots + a_0$ of degree $n \geq 1$ has precisely n roots in $\mathbb{C}$. If these roots are denoted by $w_1, \ldots, w_n$, then P can be factored as*

$$P(z) = a_n(z - w_1)(z - w_2) \cdots (z - w_n).$$

Proof. By the previous result P has a root, say w_1. Then, writing $z = (z - w_1) + w_1$, inserting this expression for z in P, and using the binomial formula we get

$$P(z) = b_n(z - w_1)^n + \cdots + b_1(z - w_1) + b_0,$$

where $b_0, \ldots, b_{n-1}$ are new coefficients, and $b_n = a_n$. Since $P(w_1) = 0$, we find that $b_0 = 0$, therefore

$$P(z) = (z - w_1)\left[b_n(z - w_1)^{n-1} + \cdots + b_1\right] = (z - w_1)Q(z),$$

where Q is a polynomial of degree $n - 1$. By induction on the degree of the polynomial, we conclude that $P(z)$ has n roots and can be expressed as

$$P(z) = c(z - w_1)(z - w_2) \cdots (z - w_n)$$

for some $c \in \mathbb{C}$. Expanding the right-hand side, we realize that the coefficient of z^n is c and therefore $c = a_n$ as claimed.

Finally, we end this section with a discussion of analytic continuation (the third of the "miracles" we mentioned in the introduction). It states that the "genetic code" of a holomorphic function is determined (that is, the function is fixed) if we know its values on appropriate arbitrarily small subsets. Note that in the theorem below, Ω is assumed connected.

Theorem 4.8 *Suppose f is a holomorphic function in a region Ω that vanishes on a sequence of distinct points with a limit point in Ω. Then f is identically* 0.

In other words, if the zeros of a holomorphic function f in the connected open set Ω accumulate *in* Ω, then $f = 0$.

Proof. Suppose that $z_0 \in \Omega$ is a limit point for the sequence $\{w_k\}_{k=1}^{\infty}$ and that $f(w_k) = 0$. First, we show that f is identically zero in a small disc containing z_0. For that, we choose a disc D centered at z_0 and contained in Ω, and consider the power series expansion of f in that disc

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

If f is not identically zero, there exists a smallest integer m such that $a_m \neq 0$. But then we can write

$$f(z) = a_m (z - z_0)^m (1 + g(z - z_0)),$$

where $g(z - z_0)$ converges to 0 as $z \to z_0$. Taking $z = w_k \neq z_0$ for a sequence of points converging to z_0, we get a contradiction since $a_m (w_k - z_0)^m \neq 0$ and $1 + g(w_k - z_0) \neq 0$, but $f(w_k) = 0$.

We conclude the proof using the fact that Ω is connected. Let U denote the interior of the set of points where $f(z) = 0$. Then U is open by definition and non-empty by the argument just given. The set U is also closed since if $z_n \in U$ and $z_n \to z$, then $f(z) = 0$ by continuity, and f vanishes in a neighborhood of z by the argument above. Hence $z \in U$. Now if we let V denote the complement of U in Ω, we conclude that U and V are both open, disjoint, and

$$\Omega = U \cup V.$$

Since Ω is connected we conclude that either U or V is empty. (Here we use one of the two equivalent definitions of connectedness discussed in Chapter 1.) Since $z_0 \in U$, we find that $U = \Omega$ and the proof is complete.

An immediate consequence of the theorem is the following.

Corollary 4.9 *Suppose f and g are holomorphic in a region Ω and $f(z) = g(z)$ for all z in some non-empty open subset of Ω (or more generally for z in some sequence of distinct points with limit point in Ω). Then $f(z) = g(z)$ throughout Ω.*

Suppose we are given a pair of functions f and F analytic in regions Ω and Ω', respectively, with $\Omega \subset \Omega'$. If the two functions agree on the smaller set Ω, we say that F is an **analytic continuation** of f into the region Ω'. The corollary then guarantees that there can be only one such analytic continuation, since F is uniquely determined by f.

5 Further applications

We gather in this section various consequences of the results proved so far.

5.1 Morera's theorem

A direct application of what was proved here is a converse of Cauchy's theorem.

Theorem 5.1 *Suppose f is a continuous function in the open disc D such that for any triangle T contained in D*

$$\int_T f(z)\,dz = 0,$$

then f is holomorphic.

Proof. By the proof of Theorem 2.1 the function f has a primitive F in D that satisfies $F' = f$. By the regularity theorem, we know that F is indefinitely (and hence twice) complex differentiable, and therefore f is holomorphic.

5.2 Sequences of holomorphic functions

Theorem 5.2 *If $\{f_n\}_{n=1}^{\infty}$ is a sequence of holomorphic functions that converges uniformly to a function f in every compact subset of Ω, then f is holomorphic in Ω.*

Proof. Let D be any disc whose closure is contained in Ω and T any triangle in that disc. Then, since each f_n is holomorphic, Goursat's theorem implies

$$\int_T f_n(z)\,dz = 0 \qquad \text{for all } n.$$

By assumption $f_n \to f$ uniformly in the closure of D, so f is continuous and

$$\int_T f_n(z)\,dz \to \int_T f(z)\,dz.$$

As a result, we find $\int_T f(z)\,dz = 0$, and by Morera's theorem, we conclude that f is holomorphic in D. Since this conclusion is true for every D whose closure is contained in Ω, we find that f is holomorphic in all of Ω.

This is a striking result that is obviously not true in the case of real variables: the uniform limit of continuously differentiable functions need not be differentiable. For example, we know that every continuous function on $[0,1]$ can be approximated uniformly by polynomials, by Weierstrass's theorem (see Chapter 5, Book I), yet not every continuous function is differentiable.

We can go one step further and deduce convergence theorems for the sequence of derivatives. Recall that if f is a power series with radius of convergence R, then f' can be obtained by differentiating term by term the series for f, and moreover f' has radius of convergence R. (See Theorem 2.6 in Chapter 1.) This implies in particular that if S_n are the partial sums of f, then S_n' converges uniformly to f' on every compact subset of the disc of convergence of f. The next theorem generalizes this fact.

Theorem 5.3 *Under the hypotheses of the previous theorem, the sequence of derivatives* $\{f_n'\}_{n=1}^{\infty}$ *converges uniformly to* f' *on every compact subset of* Ω.

Proof. We may assume without loss of generality that the sequence of functions in the theorem converges uniformly on all of Ω. Given $\delta > 0$, let Ω_δ denote the subset of Ω defined by

$$\Omega_\delta = \{z \in \Omega : \overline{D_\delta}(z) \subset \Omega\}.$$

In other words, Ω_δ consists of all points in Ω which are at distance $> \delta$ from its boundary. To prove the theorem, it suffices to show that $\{f_n'\}$ converges uniformly to f' on Ω_δ for each δ. This is achieved by proving the following inequality:

$$\sup_{z\in\Omega_\delta} |F'(z)| \leq \frac{1}{\delta} \sup_{\zeta\in\Omega} |F(\zeta)| \tag{13}$$

whenever F is holomorphic in Ω, since it can then be applied to $F = f_n - f$ to prove the desired fact. The inequality (13) follows at once from the Cauchy integral formula and the definition of Ω_δ, since for every $z \in \Omega_\delta$ the closure of $D_\delta(z)$ is contained in Ω and

$$F'(z) = \frac{1}{2\pi i} \int_{C_\delta(z)} \frac{F(\zeta)}{(\zeta - z)^2}\, d\zeta.$$

Hence,

$$\begin{aligned}|F'(z)| &\leq \frac{1}{2\pi}\int_{C_\delta(z)} \frac{|F(\zeta)|}{|\zeta - z|^2}\,|d\zeta| \\ &\leq \frac{1}{2\pi}\sup_{\zeta\in\Omega}\ |F(\zeta)|\,\frac{1}{\delta^2}\,2\pi\delta \\ &= \frac{1}{\delta}\sup_{\zeta\in\Omega}\ |F(\zeta)|\,,\end{aligned}$$

as was to be shown.

Of course, there is nothing special about the first derivative, and in fact under the hypotheses of the last theorem, we may conclude (arguing as above) that for every $k \geq 0$ the sequence of kth derivatives $\{f_n^{(k)}\}_{n=1}^\infty$ converges uniformly to $f^{(k)}$ on every compact subset of Ω.

In practice, one often uses Theorem 5.2 to construct holomorphic functions (say, with a prescribed property) as a series

$$F(z) = \sum_{n=1}^\infty f_n(z). \tag{14}$$

Indeed, if each f_n is holomorphic in a given region Ω of the complex plane, and the series converges uniformly in compact subsets of Ω, then Theorem 5.2 guarantees that F is also holomorphic in Ω. For instance, various special functions are often expressed in terms of a converging series like (14). A specific example is the Riemann zeta function discussed in Chapter 6.

We now turn to a variant of this idea, which consists of functions defined in terms of integrals.

5.3 Holomorphic functions defined in terms of integrals

As we shall see later in this book, a number of other special functions are defined in terms of integrals of the type

$$f(z) = \int_a^b F(z,s)\,ds,$$

or as limits of such integrals. Here, the function F is holomorphic in the first argument, and continuous in the second. The integral is taken in the sense of Riemann integration over the bounded interval $[a,b]$. The problem then is to establish that f is holomorphic.

In the next theorem, we impose a sufficient condition on F, often satisfied in practice, that easily implies that f is holomorphic.

After a simple linear change of variables, we may assume that $a = 0$ and $b = 1$.

Theorem 5.4 *Let $F(z,s)$ be defined for $(z,s) \in \Omega \times [0,1]$ where Ω is an open set in $\mathbb{C}$. Suppose F satisfies the following properties:*

(i) *$F(z,s)$ is holomorphic in z for each s.*

(ii) *F is continuous on $\Omega \times [0,1]$.*

Then the function f defined on Ω by

$$f(z) = \int_0^1 F(z,s)\, ds$$

is holomorphic.

The second condition says that F is jointly continuous in both arguments.

To prove this result, it suffices to prove that f is holomorphic in any disc D contained in Ω, and by Morera's theorem this could be achieved by showing that for any triangle T contained in D we have

$$\int_T \int_0^1 F(z,s)\, ds\, dz = 0.$$

Interchanging the order of integration, and using property (i) would then yield the desired result. We can, however, get around the issue of justifying the change in the order of integration by arguing differently. The idea is to interpret the integral as a "uniform" limit of Riemann sums, and then apply the results of the previous section.

Proof. For each $n \geq 1$, we consider the Riemann sum

$$f_n(z) = (1/n) \sum_{k=1}^{n} F(z, k/n).$$

Then f_n is holomorphic in all of Ω by property (i), and we claim that on any disc D whose closure is contained in Ω, the sequence $\{f_n\}_{n=1}^{\infty}$ converges uniformly to f. To see this, we recall that a continuous function on a compact set is uniformly continuous, so if $\epsilon > 0$ there exists $\delta > 0$ such that

$$\sup_{z \in D} |F(z,s_1) - F(z,s_2)| < \epsilon \qquad \text{whenever } |s_1 - s_2| < \delta.$$

Then, if $n > 1/\delta$, and $z \in D$ we have

$$\begin{aligned}|f_n(z) - f(z)| &= \left| \sum_{k=1}^{n} \int_{(k-1)/n}^{k/n} F(z, k/n) - F(z, s)\, ds \right| \\ &\leq \sum_{k=1}^{n} \int_{(k-1)/n}^{k/n} |F(z, k/n) - F(z, s)|\; ds \\ &< \sum_{k=1}^{n} \frac{\epsilon}{n} \\ &< \epsilon.\end{aligned}$$

This proves the claim, and by Theorem 5.2 we conclude that f is holomorphic in D. As a consequence, f is holomorphic in Ω, as was to be shown.

5.4 Schwarz reflection principle

In real analysis, there are various situations where one wishes to extend a function from a given set to a larger one. Several techniques exist that provide extensions for continuous functions, and more generally for functions with varying degrees of smoothness. Of course, the difficulty of the technique increases as we impose more conditions on the extension.

The situation is very different for holomorphic functions. Not only are these functions indefinitely differentiable in their domain of definition, but they also have additional characteristically rigid properties, which make them difficult to mold. For example, there exist holomorphic functions in a disc which are continuous on the closure of the disc, but which cannot be continued (analytically) into any region larger than the disc. (This phenomenon is discussed in Problem 1.) Another fact we have seen above is that holomorphic functions must be identically zero if they vanish on small open sets (or even, for example, a non-zero line segment).

It turns out that the theory developed in this chapter provides a simple extension phenomenon that is very useful in applications: the Schwarz reflection principle. The proof consists of two parts. First we define the extension, and then check that the resulting function is still holomorphic. We begin with this second point.

Let Ω be an open subset of $\mathbb{C}$ that is symmetric with respect to the real line, that is

$$z \in \Omega \quad \text{if and only if} \quad \overline{z} \in \Omega.$$

Let Ω^+ denote the part of Ω that lies in the upper half-plane and Ω^- that part that lies in the lower half-plane.

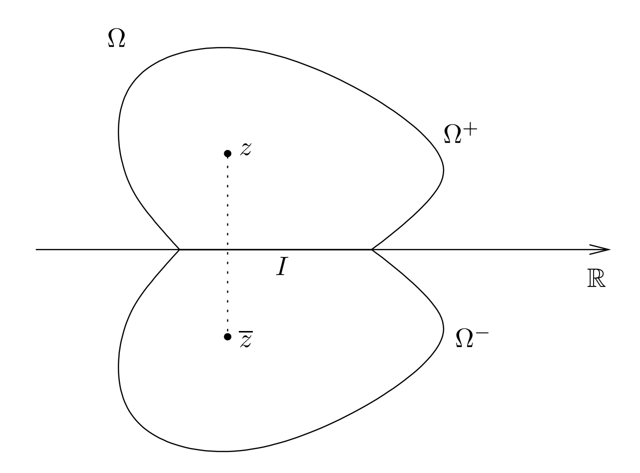


Figure 11. An open set symmetric across the real axis

Also, let $I = \Omega \cap \mathbb{R}$ so that I denotes the interior of that part of the boundary of Ω^+ and Ω^- that lies on the real axis. Then we have

$$\Omega^+ \cup I \cup \Omega^- = \Omega$$

and the only interesting case of the next theorem occurs, of course, when I is non-empty.

Theorem 5.5 (Symmetry principle) *If f^+ and f^- are holomorphic functions in Ω^+ and Ω^- respectively, that extend continuously to I and*

$$f^+(x) = f^-(x) \quad \text{for all } x \in I,$$

then the function f defined on Ω by

$$f(z) = \begin{cases} f^+(z) & \text{if } z \in \Omega^+, \\ f^+(z) = f^-(z) & \text{if } z \in I, \\ f^-(z) & \text{if } z \in \Omega^- \end{cases}$$

is holomorphic on all of Ω.

Proof. One notes first that f is continuous throughout Ω. The only difficulty is to prove that f is holomorphic at points of I. Suppose D is a

disc centered at a point on I and entirely contained in Ω. We prove that f is holomorphic in D by Morera's theorem. Suppose T is a triangle in D. If T does not intersect I, then

$$\int_T f(z)\,dz = 0$$

since f is holomorphic in the upper and lower half-discs. Suppose now that one side or vertex of T is contained in I, and the rest of T is in, say, the upper half-disc. If T_ϵ is the triangle obtained from T by slightly raising the edge or vertex which lies on I, we have $\int_{T_\epsilon} f = 0$ since T_ϵ is entirely contained in the upper half-disc (an illustration of the case when an edge lies on I is given in Figure 12(a)). We then let $\epsilon \to 0$, and by continuity we conclude that

$$\int_T f(z)\,dz = 0.$$

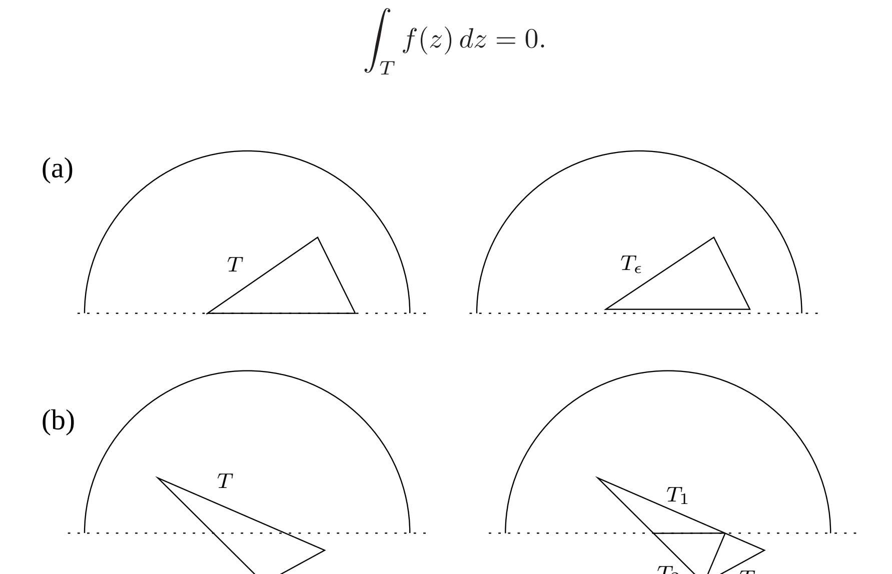


Figure 12. (a) Raising a vertex; (b) splitting a triangle

If the interior of T intersects I, we can reduce the situation to the previous one by writing T as the union of triangles each of which has an edge or vertex on I as shown in Figure 12(b). By Morera's theorem we conclude that f is holomorphic in D, as was to be shown.

We can now state the extension principle, where we use the above notation.

Theorem 5.6 (Schwarz reflection principle) *Suppose that f is a holomorphic function in Ω^+ that extends continuously to I and such that f is real-valued on I. Then there exists a function F holomorphic in all of Ω such that $F = f$ on Ω^+.*

Proof. The idea is simply to define $F(z)$ for $z \in \Omega^-$ by

$$F(z) = \overline{f(\overline{z})}.$$

To prove that F is holomorphic in Ω^- we note that if $z, z_0 \in \Omega^-$, then $\overline{z}, \overline{z_0} \in \Omega^+$ and hence, the power series expansion of f near $\overline{z_0}$ gives

$$f(\overline{z}) = \sum a_n(\overline{z} - \overline{z_0})^n.$$

As a consequence we see that

$$F(z) = \sum \overline{a_n}(z - z_0)^n$$

and F is holomorphic in Ω^-. Since f is real valued on I we have $\overline{f(x)} = f(x)$ whenever $x \in I$ and hence F extends continuously up to I. The proof is complete once we invoke the symmetry principle.

5.5 Runge's approximation theorem

We know by Weierstrass's theorem that any continuous function on a compact interval can be approximated uniformly by polynomials.[4] With this result in mind, one may inquire about similar approximations in complex analysis. More precisely, we ask the following question: what conditions on a compact set $K \subset \mathbb{C}$ guarantee that any function holomorphic in a neighborhood of this set can be approximated uniformly by polynomials on K?

An example of this is provided by power series expansions. We recall that if f is a holomorphic function in a disc D, then it has a power series expansion $f(z) = \sum_{n=0}^{\infty} a_n z^n$ that converges uniformly on every compact set $K \subset D$. By taking partial sums of this series, we conclude that f can be approximated uniformly by polynomials on any compact subset of D.

In general, however, some condition on K must be imposed, as we see by considering the function $f(z) = 1/z$ on the unit circle $K = C$. Indeed, recall that $\int_C f(z)\,dz = 2\pi i$, and if p is any polynomial, then Cauchy's theorem implies $\int_C p(z)\,dz = 0$, and this quickly leads to a contradiction.

[4] A proof may be found in Section 1.8, Chapter 5, of Book I.

A restriction on K that guarantees the approximation pertains to the topology of its complement: K^c must be connected. In fact, a slight modification of the above example when $f(z) = 1/z$ proves that this condition on K is also necessary; see Problem 4.

Conversely, uniform approximations exist when K^c is connected, and this result follows from a theorem of Runge which states that for *any* K a uniform approximation exists by *rational functions* with "singularities" in the complement of K.[5] This result is remarkable since rational functions are globally defined, while f is given only in a neighborhood of K. In particular, f could be defined independently on different components of K, making the conclusion of the theorem even more striking.

Theorem 5.7 *Any function holomorphic in a neighborhood of a compact set K can be approximated uniformly on K by rational functions whose singularities are in K^c.*

If K^c is connected, any function holomorphic in a neighborhood of K can be approximated uniformly on K by polynomials.

We shall see how the second part of the theorem follows from the first: when K^c is connected, one can "push" the singularities to infinity thereby transforming the rational functions into polynomials.

The key to the theorem lies in an integral representation formula that is a simple consequence of the Cauchy integral formula applied to a square.

Lemma 5.8 *Suppose f is holomorphic in an open set Ω, and $K \subset \Omega$ is compact. Then, there exists finitely many segments $\gamma_1, \ldots, \gamma_N$ in $\Omega - K$ such that*

$$f(z) = \sum_{n=1}^{N} \frac{1}{2\pi i} \int_{\gamma_n} \frac{f(\zeta)}{\zeta - z}\, d\zeta \quad \text{for all } z \in K. \tag{15}$$

Proof. Let $d = c \cdot d(K, \Omega^c)$, where c is any constant $< 1/\sqrt{2}$, and consider a grid formed by (solid) squares with sides parallel to the axis and of length d.

We let $\mathcal{Q} = \{Q_1, \ldots, Q_M\}$ denote the finite collection of squares in this grid that intersect K, with the boundary of each square given the positive orientation. (We denote by ∂Q_m the boundary of the square Q_m.) Finally, we let $\gamma_1, \ldots, \gamma_N$ denote the sides of squares in $\mathcal{Q}$ that do not belong to two adjacent squares in $\mathcal{Q}$. (See Figure 13.) The choice of d guarantees that for each n, $\gamma_n \subset \Omega$, and γ_n does not intersect K; for if it did, then it would belong to two adjacent squares in $\mathcal{Q}$, contradicting our choice of γ_n.

[5] These singularities are points where the function is not holomorphic, and are "poles", as defined in the next chapter.

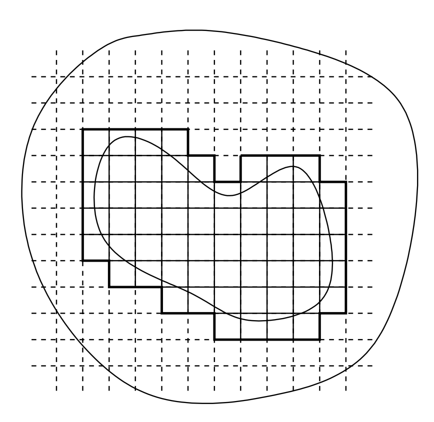


Figure 13. The union of the γ_n's is in bold-face

Since for any $z \in K$ that is not on the boundary of a square in $\mathcal{Q}$ there exists j so that $z \in Q_j$, Cauchy's theorem implies

$$\frac{1}{2\pi i}\int_{\partial Q_m} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \begin{cases} f(z) & \text{if } m = j, \\ 0 & \text{if } m \neq j. \end{cases}$$

Thus, for all such z we have

$$f(z) = \sum_{m=1}^{M} \frac{1}{2\pi i}\int_{\partial Q_m} \frac{f(\zeta)}{\zeta - z}\, d\zeta.$$

However, if Q_m and $Q_{m'}$ are adjacent, the integral over their common side is taken once in each direction, and these cancel. This establishes (15) when z is in K and not on the boundary of a square in $\mathcal{Q}$. Since $\gamma_n \subset K^c$, continuity guarantees that this relation continues to hold for all $z \in K$, as was to be shown.

The first part of Theorem 5.7 is therefore a consequence of the next lemma.

Lemma 5.9 *For any line segment γ entirely contained in $\Omega - K$, there exists a sequence of rational functions with singularities on γ that approximate the integral $\int_\gamma f(\zeta)/(\zeta - z)\, d\zeta$ uniformly on K.*

Proof. If $\gamma(t) : [0,1] \to \mathbb{C}$ is a parametrization for γ, then

$$\int_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta = \int_0^1 \frac{f(\gamma(t))}{\gamma(t) - z}\gamma'(t)\, dt.$$

Since γ does not intersect K, the integrand $F(z,t)$ in this last integral is jointly continuous on $K \times [0,1]$, and since K is compact, given $\epsilon > 0$, there exists $\delta > 0$ such that

$$\sup_{z \in K} |F(z,t_1) - F(z,t_2)| < \epsilon \quad \text{whenever } |t_1 - t_2| < \delta.$$

Arguing as in the proof of Theorem 5.4, we see that the Riemann sums of the integral $\int_0^1 F(z,t)\,dt$ approximate it uniformly on K. Since each of these Riemann sums is a rational function with singularities on γ, the lemma is proved.

Finally, the process of pushing the poles to infinity is accomplished by using the fact that K^c is connected. Since any rational function whose only singularity is at the point z_0 is a polynomial in $1/(z - z_0)$, it suffices to establish the next lemma to complete the proof of Theorem 5.7.

Lemma 5.10 *If K^c is connected and $z_0 \notin K$, then the function $1/(z - z_0)$ can be approximated uniformly on K by polynomials.*

Proof. First, we choose a point z_1 that is outside a large open disc D centered at the origin and which contains K. Then

$$\frac{1}{z - z_1} = -\frac{1}{z_1}\frac{1}{1 - z/z_1} = \sum_{n=1}^{\infty} -\frac{z^n}{z_1^{n+1}},$$

where the series converges uniformly for $z \in K$. The partial sums of this series are polynomials that provide a uniform approximation to $1/(z - z_1)$ on K. In particular, this implies that any power $1/(z - z_1)^k$ can also be approximated uniformly on K by polynomials.

It now suffices to prove that $1/(z - z_0)$ can be approximated uniformly on K by polynomials in $1/(z - z_1)$. To do so, we use the fact that K^c is connected to travel from z_0 to the point z_1. Let γ be a curve in K^c that is parametrized by $\gamma(t)$ on $[0,1]$, and such that $\gamma(0) = z_0$ and $\gamma(1) = z_1$. If we let $\rho = \frac{1}{2}d(K,\gamma)$, then $\rho > 0$ since γ and K are compact. We then choose a sequence of points $\{w_1, \ldots, w_\ell\}$ on γ such that $w_0 = z_0$, $w_\ell = z_1$, and $|w_j - w_{j+1}| < \rho$ for all $0 \leq j < \ell$.

We claim that if w is a point on γ, and w' any other point with $|w - w'| < \rho$, then $1/(z - w)$ can be approximated uniformly on K by polynomials in $1/(z - w')$. To see this, note that

$$\begin{aligned}\frac{1}{z - w} &= \frac{1}{z - w'}\frac{1}{1 - \frac{w - w'}{z - w'}} \\ &= \sum_{n=0}^{\infty} \frac{(w - w')^n}{(z - w')^{n+1}},\end{aligned}$$

and since the sum converges uniformly for $z \in K$, the approximation by partial sums proves our claim.

This result allows us to travel from z_0 to z_1 through the finite sequence $\{w_j\}$ to find that $1/(z - z_0)$ can be approximated uniformly on K by polynomials in $1/(z - z_1)$. This concludes the proof of the lemma, and also that of the theorem.

6 Exercises

1. Prove that

$$\int_0^\infty \sin(x^2)\,dx = \int_0^\infty \cos(x^2)\,dx = \frac{\sqrt{2\pi}}{4}.$$

These are the **Fresnel integrals**. Here, $\int_0^\infty$ is interpreted as $\lim_{R\to\infty} \int_0^R$.

[Hint: Integrate the function e^{-z^2} over the path in Figure 14. Recall that $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$.]

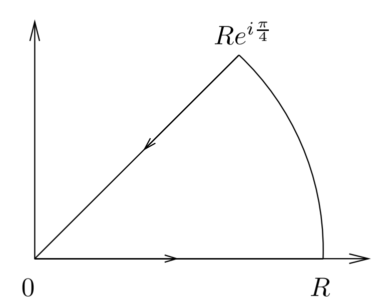


Figure 14. The contour in Exercise 1

2. Show that $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$.

[Hint: The integral equals $\frac{1}{2i}\int_{-\infty}^{\infty} \frac{e^{ix}-1}{x}\,dx$. Use the indented semicircle.]

3. Evaluate the integrals

$$\int_0^\infty e^{-ax}\cos bx\,dx \quad \text{and} \quad \int_0^\infty e^{-ax}\sin bx\,dx\,, \qquad a > 0$$

by integrating e^{-Az}, $A = \sqrt{a^2 + b^2}$, over an appropriate sector with angle ω, with $\cos\omega = a/A$.

4. Prove that for all $\xi \in \mathbb{C}$ we have $e^{-\pi\xi^2} = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{2\pi i x\xi}\, dx.$

5. Suppose f is continuously *complex* differentiable on Ω, and $T \subset \Omega$ is a triangle whose interior is also contained in Ω. Apply Green's theorem to show that

$$\int_T f(z)\, dz = 0.$$

This provides a proof of Goursat's theorem under the additional assumption that f' is continuous.

[Hint: Green's theorem says that if (F, G) is a continuously differentiable vector field, then

$$\int_T F\, dx + G\, dy = \int_{\text{Interior of } T} \left(\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} \right) dxdy.$$

For appropriate F and G, one can then use the Cauchy-Riemann equations.]

6. Let Ω be an open subset of $\mathbb{C}$ and let $T \subset \Omega$ be a triangle whose interior is also contained in Ω. Suppose that f is a function holomorphic in Ω except possibly at a point w inside T. Prove that if f is bounded near w, then

$$\int_T f(z)\, dz = 0.$$

7. Suppose $f : \mathbb{D} \to \mathbb{C}$ is holomorphic. Show that the diameter $d = \sup_{z,\, w \in \mathbb{D}} |f(z) - f(w)|$ of the image of f satisfies

$$2|f'(0)| \leq d.$$

Moreover, it can be shown that equality holds precisely when f is linear, $f(z) = a_0 + a_1 z$.

Note. In connection with this result, see the relationship between the diameter of a curve and Fourier series described in Problem 1, Chapter 4, Book I.

[Hint: $2f'(0) = \frac{1}{2\pi i} \int_{|\zeta|=r} \frac{f(\zeta) - f(-\zeta)}{\zeta^2}\, d\zeta$ whenever $0 < r < 1$.]

8. If f is a holomorphic function on the strip $-1 < y < 1$, $x \in \mathbb{R}$ with

$$|f(z)| \leq A(1 + |z|)^\eta, \qquad \eta \text{ a fixed real number}$$

for all z in that strip, show that for each integer $n \geq 0$ there exists $A_n \geq 0$ so that

$$|f^{(n)}(x)| \leq A_n(1 + |x|)^\eta, \qquad \text{for all } x \in \mathbb{R}.$$

[Hint: Use the Cauchy inequalities.]

9. Let Ω be a bounded open subset of $\mathbb{C}$, and $\varphi : \Omega \to \Omega$ a holomorphic function. Prove that if there exists a point $z_0 \in \Omega$ such that

$$\varphi(z_0) = z_0 \quad \text{and} \quad \varphi'(z_0) = 1$$

then φ is linear.

[Hint: Why can one assume that $z_0 = 0$? Write $\varphi(z) = z + a_n z^n + O(z^{n+1})$ near 0, and prove that if $\varphi_k = \varphi \circ \cdots \circ \varphi$ (where φ appears k times), then $\varphi_k(z) = z + ka_n z^n + O(z^{n+1})$. Apply the Cauchy inequalities and let $k \to \infty$ to conclude the proof. Here we use the standard O notation, where $f(z) = O(g(z))$ as $z \to 0$ means that $|f(z)| \leq C|g(z)|$ for some constant C as $|z| \to 0$.]

10. Weierstrass's theorem states that a continuous function on $[0, 1]$ can be uniformly approximated by polynomials. Can every continuous function on the closed unit disc be approximated uniformly by polynomials in the variable z?

11. Let f be a holomorphic function on the disc D_{R_0} centered at the origin and of radius R_0.

(a) Prove that whenever $0 < R < R_0$ and $|z| < R$, then

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(Re^{i\varphi}) \text{Re} \left(\frac{Re^{i\varphi} + z}{Re^{i\varphi} - z} \right) d\varphi.$$

(b) Show that

$$\text{Re} \left(\frac{Re^{i\gamma} + r}{Re^{i\gamma} - r} \right) = \frac{R^2 - r^2}{R^2 - 2Rr\cos\gamma + r^2}.$$

[Hint: For the first part, note that if $w = R^2/\overline{z}$, then the integral of $f(\zeta)/(\zeta - w)$ around the circle of radius R centered at the origin is zero. Use this, together with the usual Cauchy integral formula, to deduce the desired identity.]

12. Let u be a real-valued function defined on the unit disc $\mathbb{D}$. Suppose that u is twice continuously differentiable and harmonic, that is,

$$\triangle u(x, y) = 0$$

for all $(x, y) \in \mathbb{D}$.

(a) Prove that there exists a holomorphic function f on the unit disc such that

$$\text{Re}(f) = u.$$

Also show that the imaginary part of f is uniquely defined up to an additive (real) constant. [Hint: From the previous chapter we would have $f'(z) = 2\partial u/\partial z$. Therefore, let $g(z) = 2\partial u/\partial z$ and prove that g is holomorphic. Why can one find F with $F' = g$? Prove that $\text{Re}(F)$ differs from u by a real constant.]

(b) Deduce from this result, and from Exercise 11, the Poisson integral representation formula from the Cauchy integral formula: If u is harmonic in the unit disc and continuous on its closure, then if $z = re^{i\theta}$ one has

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} P_r(\theta - \varphi) u(\varphi)\, d\varphi$$

where $P_r(\gamma)$ is the Poisson kernel for the unit disc given by

$$P_r(\gamma) = \frac{1 - r^2}{1 - 2r\cos\gamma + r^2}.$$

13. Suppose f is an analytic function defined everywhere in $\mathbb{C}$ and such that for each $z_0 \in \mathbb{C}$ at least one coefficient in the expansion

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n$$

is equal to 0. Prove that f is a polynomial.
[Hint: Use the fact that $c_n n! = f^{(n)}(z_0)$ and use a countability argument.]

14. Suppose that f is holomorphic in an open set containing the closed unit disc, except for a pole at z_0 on the unit circle. Show that if

$$\sum_{n=0}^{\infty} a_n z^n$$

denotes the power series expansion of f in the open unit disc, then

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = z_0.$$

15. Suppose f is a non-vanishing continuous function on $\overline{\mathbb{D}}$ that is holomorphic in $\mathbb{D}$. Prove that if

$$|f(z)| = 1 \qquad \text{whenever } |z| = 1,$$

then f is constant.
[Hint: Extend f to all of $\mathbb{C}$ by $f(z) = 1/\overline{f(1/\overline{z})}$ whenever $|z| > 1$, and argue as in the Schwarz reflection principle.]

7 Problems

1. Here are some examples of analytic functions on the unit disc that cannot be extended analytically past the unit circle. The following definition is needed. Let

f be a function defined in the unit disc $\mathbb{D}$, with boundary circle C. A point w on C is said to be *regular* for f if there is an open neighborhood U of w and an analytic function g on U, so that $f = g$ on $\mathbb{D} \cap U$. A function f defined on $\mathbb{D}$ cannot be continued analytically past the unit circle if no point of C is regular for f.

(a) Let

$$f(z) = \sum_{n=0}^{\infty} z^{2^n} \qquad \text{for } |z| < 1.$$

Notice that the radius of convergence of the above series is 1. Show that f cannot be continued analytically past the unit disc. [Hint: Suppose $\theta = 2\pi p/2^k$, where p and k are positive integers. Let $z = re^{i\theta}$; then $|f(re^{i\theta})| \to \infty$ as $r \to 1$.]

(b) * Fix $0 < \alpha < \infty$. Show that the analytic function f defined by

$$f(z) = \sum_{n=0}^{\infty} 2^{-n\alpha} z^{2^n} \qquad \text{for } |z| < 1$$

extends continuously to the unit circle, but cannot be analytically continued past the unit circle. [Hint: There is a nowhere differentiable function lurking in the background. See Chapter 4 in Book I.]

2.* Let

$$F(z) = \sum_{n=1}^{\infty} d(n) z^n \qquad \text{for } |z| < 1$$

where $d(n)$ denotes the number of divisors of n. Observe that the radius of convergence of this series is 1. Verify the identity

$$\sum_{n=1}^{\infty} d(n) z^n = \sum_{n=1}^{\infty} \frac{z^n}{1 - z^n}.$$

Using this identity, show that if $z = r$ with $0 < r < 1$, then

$$|F(r)| \geq c \frac{1}{1-r} \log(1/(1-r))$$

as $r \to 1$. Similarly, if $\theta = 2\pi p/q$ where p and q are positive integers and $z = re^{i\theta}$, then

$$|F(re^{i\theta})| \geq c_{p/q} \frac{1}{1-r} \log(1/(1-r))$$

as $r \to 1$. Conclude that F cannot be continued analytically past the unit disc.

3. Morera's theorem states that if f is continuous in $\mathbb{C}$, and $\int_T f(z)\, dz = 0$ for all triangles T, then f is holomorphic in $\mathbb{C}$. Naturally, we may ask if the conclusion still holds if we replace triangles by other sets.

(a) Suppose that f is continuous on $\mathbb{C}$, and

$$\int_C f(z)\,dz = 0 \tag{16}$$

for every circle C. Prove that f is holomorphic.

(b) More generally, let Γ be any toy contour, and $\mathcal{F}$ the collection of all translates and dilates of Γ. Show that if f is continuous on $\mathbb{C}$, and

$$\int_\gamma f(z)\,dz = 0 \qquad \text{for all } \gamma \in \mathcal{F}$$

then f is holomorphic. In particular, Morera's theorem holds under the weaker assumption that $\int_T f(z)\,dz = 0$ for all equilateral triangles.

[Hint: As a first step, assume that f is twice real differentiable, and write $f(z) = f(z_0) + a(z - z_0) + b(\overline{z - z_0}) + O(|z - z_0|^2)$ for z near z_0. Integrating this expansion over small circles around z_0 yields $\partial f/\partial\overline{z} = b = 0$ at z_0. Alternatively, suppose only that f is differentiable and apply Green's theorem to conclude that the real and imaginary parts of f satisfy the Cauchy-Riemann equations.

In general, let $\varphi(w) = \varphi(x, y)$ (when $w = x + iy$) denote a smooth function with $0 \leq \varphi(w) \leq 1$, and $\int_{\mathbb{R}^2} \varphi(w)\,dV(w) = 1$, where $dV(w) = dxdy$, and $\int$ denotes the usual integral of a function of two variables in $\mathbb{R}^2$. For each $\epsilon > 0$, let $\varphi_\epsilon(z) = \epsilon^{-2}\varphi(\epsilon^{-1}z)$, as well as

$$f_\epsilon(z) = \int_{\mathbb{R}^2} f(z - w)\varphi_\epsilon(w)\,dV(w),$$

where the integral denotes the usual integral of functions of two variables, with $dV(w)$ the area element of $\mathbb{R}^2$. Then f_ϵ is smooth, satisfies condition (16), and $f_\epsilon \to f$ uniformly on any compact subset of $\mathbb{C}$.]

4. Prove the converse to Runge's theorem: if K is a compact set whose complement if not connected, then there exists a function f holomorphic in a neighborhood of K which cannot be approximated uniformly by polynomial on K.

[Hint: Pick a point z_0 in a bounded component of K^c, and let $f(z) = 1/(z - z_0)$. If f can be approximated uniformly by polynomials on K, show that there exists a polynomial p such that $|(z - z_0)p(z) - 1| < 1$. Use the maximum modulus principle (Chapter 3) to show that this inequality continues to hold for all z in the component of K^c that contains z_0.]

5.* There exists an entire function F with the following "universal" property: given any entire function h, there is an increasing sequence $\{N_k\}_{k=1}^\infty$ of positive integers, so that

$$\lim_{n\to\infty} F(z + N_k) = h(z)$$

uniformly on every compact subset of $\mathbb{C}$.

(a) Let $p_1, p_2, \ldots$ denote an enumeration of the collection of polynomials whose coefficients have rational real and imaginary parts. Show that it suffices to find an entire function F and an increasing sequence $\{M_n\}$ of positive integers, such that

$$|F(z) - p_n(z - M_n)| < \frac{1}{n} \quad \text{whenever } z \in D_n, \tag{17}$$

where D_n denotes the disc centered at M_n and of radius n. [Hint: Given h entire, there exists a sequence $\{n_k\}$ such that $\lim_{k\to\infty} p_{n_k}(z) = h(z)$ uniformly on every compact subset of $\mathbb{C}$.]

(b) Construct F satisfying (17) as an infinite series

$$F(z) = \sum_{n=1}^{\infty} u_n(z)$$

where $u_n(z) = p_n(z - M_n)e^{-c_n(z-M_n)^2}$, and the quantities $c_n > 0$ and $M_n > 0$ are chosen appropriately with $c_n \to 0$ and $M_n \to \infty$. [Hint: The function e^{-z^2} vanishes rapidly as $|z| \to \infty$ in the sectors $\{|\arg z| < \pi/4 - \delta\}$ and $\{|\pi - \arg z| < \pi/4 - \delta\}$.]

In the same spirit, there exists an alternate "universal" entire function G with the following property: given any entire function h, there is an increasing sequence $\{N_k\}_{k=1}^{\infty}$ of positive integers, so that

$$\lim_{k\to\infty} D^{N_k} G(z) = h(z)$$

uniformly on every compact subset of $\mathbb{C}$. Here $D^j G$ denotes the jth (complex) derivative of G.

3 Meromorphic Functions and the Logarithm

> One knows that the differential calculus, which has contributed so much to the progress of analysis, is founded on the consideration of differential coefficients, that is derivatives of functions. When one attributes an infinitesimal increase ϵ to the variable x, the function $f(x)$ of this variable undergoes in general an infinitesimal increase of which the first term is proportional to ϵ, and the finite coefficient of ϵ of this increase is what is called its differential coefficient... If considering the values of x where $f(x)$ becomes infinite, we add to one of these values designated by x_1, the infinitesimal ϵ, and then develop $f(x_1+\epsilon)$ in increasing power of the same quantity, the first terms of this development contain negative powers of ϵ; one of these will be the product of $1/\epsilon$ with a finite coefficient, which we will call the residue of the function $f(x)$, relative to the particular value x_1 of the variable x. Residues of this kind present themselves naturally in several branches of algebraic and infinitesimal analysis. Their consideration furnish methods that can be simply used, that apply to a large number of diverse questions, and that give new formulae that would seem to be of interest to mathematicians...
>
> *A. L. Cauchy,* 1826

There is a general principle in the theory, already implicit in Riemann's work, which states that analytic functions are in an essential way characterized by their singularities. That is to say, globally analytic functions are "effectively" determined by their zeros, and meromorphic functions by their zeros and poles. While these assertions cannot be formulated as precise general theorems, there are nevertheless significant instances where this principle applies.

We begin this chapter by considering singularities, in particular the different kind of point singularities ("isolated" singularities) that a holomorphic function can have. In order of increasing severity, these are:

- removable singularities

- poles
- essential singularities.

The first type is harmless since a function can actually be extended to be holomorphic at its removable singularities (hence the name). Near the third type, the function oscillates and may grow faster than any power, and a complete understanding of its behavior is not easy. For the second type the analysis is more straight forward and is connected with the calculus of residues, which arises as follows.

Recall that by Cauchy's theorem a holomorphic function f in an open set which contains a closed curve γ and its interior satisfies

$$\int_\gamma f(z)\, dz = 0.$$

The question that occurs is: what happens if f has a pole in the interior of the curve? To try to answer this question consider the example $f(z) = 1/z$, and recall that if C is a (positively oriented) circle centered at 0, then

$$\int_C \frac{dz}{z} = 2\pi i.$$

This turns out to be the key ingredient in the calculus of residues.

A new aspect appears when we consider indefinite integrals of holomorphic functions that have singularities. As the basic example $f(z) = 1/z$ shows, the resulting "function" (in this case the logarithm) may not be single-valued, and understanding this phenomenon is of importance for a number of subjects. Exploiting this multi-valuedness leads in effect to the "argument principle." We can use this principle to count the number of zeros of a holomorphic function inside a suitable curve. As a simple consequence of this result, we obtain a significant geometric property of holomorphic functions: they are open mappings. From this, the maximum principle, another important feature of holomorphic functions, is an easy step.

In order to turn to the logarithm itself, and come to grips with the precise nature of its multi-valuedness, we introduce the notions of homotopy of curves and simply connected domains. It is on the latter type of open sets that single-valued branches of the logarithm can be defined.

1 Zeros and poles

By definition, a **point singularity** of a function f is a complex number z_0 such that f is defined in a neighborhood of z_0 but not at the point

z_0 itself. We shall also call such points **isolated singularities**. For example, if the function f is defined only on the punctured plane by $f(z) = z$, then the origin is a point singularity. Of course, in that case, the function f can actually be defined at 0 by setting $f(0) = 0$, so that the resulting extension is continuous and in fact entire. (Such points are then called removable singularities.) More interesting is the case of the function $g(z) = 1/z$ defined in the punctured plane. It is clear now that g cannot be defined as a continuous function, much less as a holomorphic function, at the point 0. In fact, $g(z)$ grows to infinity as z approaches 0, and we shall say that the origin is a pole singularity. Finally, the case of the function $h(z) = e^{1/z}$ on the punctured plane shows that removable singularities and poles do not tell the whole story. Indeed, the function $h(z)$ grows indefinitely as z approaches 0 on the positive real line, while h approaches 0 as z goes to 0 on the negative real axis. Finally h oscillates rapidly, yet remains bounded, as z approaches the origin on the imaginary axis.

Since singularities often appear because the denominator of a fraction vanishes, we begin with a local study of the zeros of a holomorphic function.

A complex number z_0 is a **zero** for the holomorphic function f if $f(z_0) = 0$. In particular, analytic continuation shows that the zeros of a non-trivial holomorphic function are isolated. In other words, if f is holomorphic in Ω and $f(z_0) = 0$ for some $z_0 \in \Omega$, then there exists an open neighborhood U of z_0 such that $f(z) \neq 0$ for all $z \in U - \{z_0\}$, unless f is identically zero. We start with a local description of a holomorphic function near a zero.

Theorem 1.1 *Suppose that f is holomorphic in a connected open set Ω, has a zero at a point $z_0 \in \Omega$, and does not vanish identically in Ω. Then there exists a neighborhood $U \subset \Omega$ of z_0, a non-vanishing holomorphic function g on U, and a unique positive integer n such that*

$$f(z) = (z - z_0)^n g(z) \quad \text{for all } z \in U.$$

Proof. Since Ω is connected and f is not identically zero, we conclude that f is not identically zero in a neighborhood of z_0. In a small disc centered at z_0 the function f has a power series expansion

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k.$$

Since f is not identically zero near z_0, there exists a smallest integer n

such that $a_n \neq 0$. Then, we can write

$$f(z) = (z - z_0)^n [a_n + a_{n+1}(z - z_0) + \cdots] = (z - z_0)^n g(z) ,$$

where g is defined by the series in brackets, and hence is holomorphic, and is nowhere vanishing for all z close to z_0 (since $a_n \neq 0$). To prove the uniqueness of the integer n, suppose that we can also write

$$f(z) = (z - z_0)^n g(z) = (z - z_0)^m h(z)$$

where $h(z_0) \neq 0$. If $m > n$, then we may divide by $(z - z_0)^n$ to see that

$$g(z) = (z - z_0)^{m-n} h(z)$$

and letting $z \to z_0$ yields $g(z_0) = 0$, a contradiction. If $m < n$ a similar argument gives $h(z_0) = 0$, which is also a contradiction. We conclude that $m = n$, thus $h = g$, and the theorem is proved.

In the case of the above theorem, we say that f has a **zero of order** n (or **multiplicity** n) at z_0. If a zero is of order 1, we say that it is **simple**. We observe that, quantitatively, the order describes the rate at which the function vanishes.

The importance of the previous theorem comes from the fact that we can now describe precisely the type of singularity possessed by the function $1/f$ at z_0.

For this purpose, it is now convenient to define a **deleted neighborhood** of z_0 to be an open disc centered at z_0, minus the point z_0, that is, the set

$$\{z : 0 < |z - z_0| < r\}$$

for some $r > 0$. Then, we say that a function f defined in a deleted neighborhood of z_0 has a **pole** at z_0, if the function $1/f$, defined to be zero at z_0, is holomorphic in a full neighborhood of z_0.

Theorem 1.2 *If f has a pole at $z_0 \in \Omega$, then in a neighborhood of that point there exist a non-vanishing holomorphic function h and a unique positive integer n such that*

$$f(z) = (z - z_0)^{-n} h(z).$$

Proof. By the previous theorem we have $1/f(z) = (z - z_0)^n g(z)$, where g is holomorphic and non-vanishing in a neighborhood of z_0, so the result follows with $h(z) = 1/g(z)$.

The integer n is called the **order** (or **multiplicity**) of the pole, and describes the rate at which the function grows near z_0. If the pole is of order 1, we say that it is **simple**.

The next theorem should be reminiscent of power series expansion, except that now we allow terms of negative order, to account for the presence of a pole.

Theorem 1.3 *If f has a pole of order n at z_0, then*

$$f(z) = \frac{a_{-n}}{(z-z_0)^n} + \frac{a_{-n+1}}{(z-z_0)^{n-1}} + \cdots + \frac{a_{-1}}{(z-z_0)} + G(z)\,, \tag{1}$$

where G is a holomorphic function in a neighborhood of z_0.

Proof. The proof follows from the multiplicative statement in the previous theorem. Indeed, the function h has a power series expansion

$$h(z) = A_0 + A_1(z-z_0) + \cdots$$

so that

$$\begin{aligned} f(z) &= (z-z_0)^{-n}(A_0 + A_1(z-z_0) + \cdots) \\ &= \frac{a_{-n}}{(z-z_0)^n} + \frac{a_{-n+1}}{(z-z_0)^{n-1}} + \cdots + \frac{a_{-1}}{(z-z_0)} + G(z). \end{aligned}$$

The sum

$$\frac{a_{-n}}{(z-z_0)^n} + \frac{a_{-n+1}}{(z-z_0)^{n-1}} + \cdots + \frac{a_{-1}}{(z-z_0)}$$

is called the **principal part** of f at the pole z_0, and the coefficient a_{-1} is the **residue** of f at that pole. We write $\text{res}_{z_0} f = a_{-1}$. The importance of the residue comes from the fact that all the other terms in the principal part, that is, those of order strictly greater than 1, have primitives in a deleted neighborhood of z_0. Therefore, if $P(z)$ denotes the principal part above and C is any circle centered at z_0, we get

$$\frac{1}{2\pi i}\int_C P(z)\,dz = a_{-1}.$$

We shall return to this important point in the section on the residue formula.

As we shall see, in many cases, the evaluation of integrals reduces to the calculation of residues. In the case when f has a simple pole at z_0, it is clear that

$$\text{res}_{z_0} f = \lim_{z\to z_0} (z-z_0)f(z).$$

If the pole is of higher order, a similar formula holds, one that involves differentiation as well as taking a limit.

Theorem 1.4 *If f has a pole of order n at z_0, then*

$$\operatorname{res}_{z_0} f = \lim_{z \to z_0} \frac{1}{(n-1)!} \left(\frac{d}{dz}\right)^{n-1} (z - z_0)^n f(z).$$

The theorem is an immediate consequence of formula (1), which implies

$$(z - z_0)^n f(z) = a_{-n} + a_{-n+1}(z - z_0) + \cdots + a_{-1}(z - z_0)^{n-1} +$$

$$+G(z)(z - z_0)^n.$$

2 The residue formula

We now discuss the celebrated residue formula. Our approach follows the discussion of Cauchy's theorem in the last chapter: we first consider the case of the circle and its interior the disc, and then explain generalizations to toy contours and their interiors.

Theorem 2.1 *Suppose that f is holomorphic in an open set containing a circle C and its interior, except for a pole at z_0 inside C. Then*

$$\int_C f(z)\, dz = 2\pi i \operatorname{res}_{z_0} f.$$

Proof. Once again, we may choose a keyhole contour that avoids the pole, and let the width of the corridor go to zero to see that

$$\int_C f(z)\, dz = \int_{C_\epsilon} f(z)\, dz$$

where C_ϵ is the small circle centered at the pole z_0 and of radius ϵ.

Now we observe that

$$\frac{1}{2\pi i} \int_{C_\epsilon} \frac{a_{-1}}{z - z_0}\, dz = a_{-1}$$

is an immediate consequence of the Cauchy integral formula (Theorem 4.1 of the previous chapter), applied to the constant function $f = a_{-1}$. Similarly,

$$\frac{1}{2\pi i} \int_{C_\epsilon} \frac{a_{-k}}{(z - z_0)^k}\, dz = 0$$

when $k > 1$, by using the corresponding formulae for the derivatives (Corollary 4.2 also in the previous chapter). But we know that in a neighborhood of z_0 we can write

$$f(z) = \frac{a_{-n}}{(z-z_0)^n} + \frac{a_{-n+1}}{(z-z_0)^{n-1}} + \cdots + \frac{a_{-1}}{z-z_0} + G(z),$$

where G is holomorphic. By Cauchy's theorem, we also know that $\int_{C_\epsilon} G(z)\,dz = 0$, hence $\int_{C_\epsilon} f(z)\,dz = a_{-1}$. This implies the desired result.

This theorem can be generalized to the case of finitely many poles in the circle, as well as to the case of toy contours.

Corollary 2.2 *Suppose that f is holomorphic in an open set containing a circle C and its interior, except for poles at the points $z_1, \ldots, z_N$ inside C. Then*

$$\int_C f(z)\,dz = 2\pi i \sum_{k=1}^{N} \operatorname{res}_{z_k} f.$$

For the proof, consider a multiple keyhole which has a loop avoiding each one of the poles. Let the width of the corridors go to zero. In the limit, the integral over the large circle equals a sum of integrals over small circles to which Theorem 2.1 applies.

Corollary 2.3 *Suppose that f is holomorphic in an open set containing a toy contour γ and its interior, except for poles at the points $z_1, \ldots, z_N$ inside γ. Then*

$$\int_\gamma f(z)\,dz = 2\pi i \sum_{k=1}^{N} \operatorname{res}_{z_k} f.$$

In the above, we take γ to have positive orientation.

The proof consists of choosing a keyhole appropriate for the given toy contour, so that, as we have seen previously, we can reduce the situation to integrating over small circles around the poles where Theorem 2.1 applies.

The identity $\int_\gamma f(z)\,dz = 2\pi i \sum_{k=1}^{N} \operatorname{res}_{z_k} f$ is referred to as the **residue formula**.

2.1 Examples

The calculus of residues provides a powerful technique to compute a wide range of integrals. In the examples we give next, we evaluate three

improper Riemann integrals of the form

$$\int_{-\infty}^{\infty} f(x)\, dx.$$

The main idea is to extend f to the complex plane, and then choose a family γ_R of toy contours so that

$$\lim_{R\to\infty} \int_{\gamma_R} f(z)\, dz = \int_{-\infty}^{\infty} f(x)\, dx.$$

By computing the residues of f at its poles, we easily obtain $\int_{\gamma_R} f(z)\, dz$. The challenging part is to choose the contours γ_R, so that the above limit holds. Often, this choice is motivated by the decay behavior of f.

EXAMPLE 1. First, we prove that

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi \tag{2}$$

by using contour integration. Note that if we make the change of variables $x \mapsto x/y$, this yields

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y\, dx}{y^2+x^2} = \int_{-\infty}^{\infty} \mathcal{P}_y(x)\, dx.$$

In other words, formula (2) says that the integral of the Poisson kernel $\mathcal{P}_y(x)$ is equal to 1 for each $y > 0$. This was proved quite easily in Lemma 2.5 of Chapter 5 in Book I, since $1/(1+x^2)$ is the derivative of $\arctan x$. Here we provide a residue calculation that leads to another proof of (2).

Consider the function

$$f(z) = \frac{1}{1+z^2}\,,$$

which is holomorphic in the complex plane except for simple poles at the points i and $-i$. Also, we choose the contour γ_R shown in Figure 1. The contour consists of the segment $[-R, R]$ on the real axis and of a large half-circle centered at the origin in the upper half-plane.
Since we may write

$$f(z) = \frac{1}{(z-i)(z+i)}$$

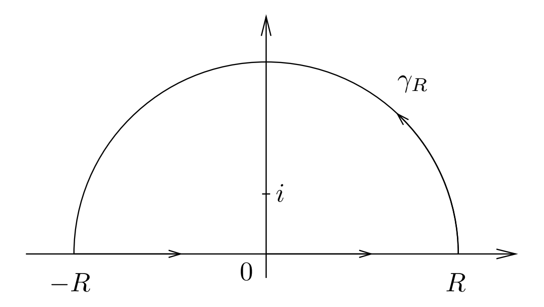


Figure 1. The contour γ_R in Example 1

we see that the residue of f at i is simply $1/2i$. Therefore, if R is large enough, we have

$$\int_{\gamma_R} f(z)\,dz = \frac{2\pi i}{2i} = \pi.$$

If we denote by C_R^+ the large half-circle of radius R, we see that

$$\left|\int_{C_R^+} f(z)\,dz\right| \leq \pi R \frac{B}{R^2} \leq \frac{M}{R},$$

where we have used the fact that $|f(z)| \leq B/|z|^2$ when $z \in C_R^+$ and R is large. So this integral goes to 0 as $R \to \infty$. Therefore, in the limit we find that

$$\int_{-\infty}^{\infty} f(x)\,dx = \pi,$$

as desired. We remark that in this example, there is nothing special about our choice of the semicircle in the upper half-plane. One gets the same conclusion if one uses the semicircle in the lower half-plane, with the other pole and the appropriate residue.

EXAMPLE 2. An integral that will play an important role in Chapter 6 is

$$\int_{-\infty}^{\infty} \frac{e^{ax}}{1+e^x}\,dx = \frac{\pi}{\sin \pi a}, \qquad 0 < a < 1.$$

To prove this formula, let $f(z) = e^{az}/(1+e^z)$, and consider the contour consisting of a rectangle in the upper half-plane with a side lying

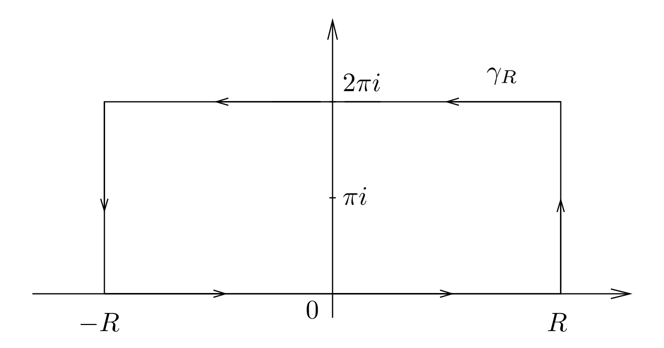


Figure 2. The contour γ_R in Example 2

on the real axis, and a parallel side on the line $\mathrm{Im}(z) = 2\pi$, as shown in Figure 2.

The only point in the rectangle γ_R where the denominator of f vanishes is $z = \pi i$. To compute the residue of f at that point, we argue as follows: First, note

$$(z - \pi i) f(z) = e^{az} \frac{z - \pi i}{1 + e^z} = e^{az} \frac{z - \pi i}{e^z - e^{\pi i}}.$$

We recognize on the right the inverse of a difference quotient, and in fact

$$\lim_{z \to \pi i} \frac{e^z - e^{\pi i}}{z - \pi i} = e^{\pi i} = -1$$

since e^z is its own derivative. Therefore, the function f has a simple pole at πi with residue

$$\mathrm{res}_{\pi i} f = -e^{a\pi i}.$$

As a consequence, the residue formula says that

$$\int_{\gamma_R} f = -2\pi i e^{a\pi i}. \tag{3}$$

We now investigate the integrals of f over each side of the rectangle. Let I_R denote

$$\int_{-R}^{R} f(x)\, dx$$

and I the integral we wish to compute, so that $I_R \to I$ as $R \to \infty$. Then, it is clear that the integral of f over the top side of the rectangle (with

the orientation from right to left) is

$$-e^{2\pi ia}I_R.$$

Finally, if $A_R = \{R + it : \ 0 \le t \le 2\pi\}$ denotes the vertical side on the right, then

$$\left|\int_{A_R} f\right| \le \int_0^{2\pi} \left|\frac{e^{a(R+it)}}{1 + e^{R+it}}\right| dt \le Ce^{(a-1)R},$$

and since $a < 1$, this integral tends to 0 as $R \to \infty$. Similarly, the integral over the vertical segment on the left goes to 0, since it can be bounded by Ce^{-aR} and $a > 0$. Therefore, in the limit as R tends to infinity, the identity (3) yields

$$I - e^{2\pi ia}I = -2\pi i e^{a\pi i},$$

from which we deduce

$$\begin{aligned} I &= -2\pi i \frac{e^{a\pi i}}{1 - e^{2\pi ia}} \\ &= \frac{2\pi i}{e^{\pi ia} - e^{-\pi ia}} \\ &= \frac{\pi}{\sin \pi a}, \end{aligned}$$

and the computation is complete.

EXAMPLE 3. Now we calculate another Fourier transform, namely

$$\int_{-\infty}^{\infty} \frac{e^{-2\pi ix\xi}}{\cosh \pi x}\, dx = \frac{1}{\cosh \pi\xi}$$

where

$$\cosh z = \frac{e^z + e^{-z}}{2}.$$

In other words, the function $1/\cosh \pi x$ is its own Fourier transform, a property also shared by $e^{-\pi x^2}$ (see Example 1, Chapter 2). To see this, we use a rectangle γ_R as shown on Figure 3 whose width goes to infinity, but whose height is fixed.

For a fixed $\xi \in \mathbb{R}$, let

$$f(z) = \frac{e^{-2\pi iz\xi}}{\cosh \pi z}$$

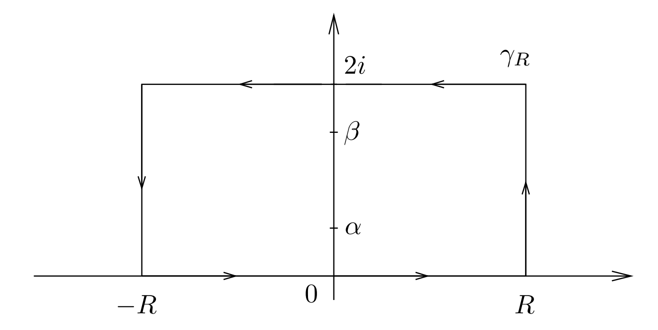


Figure 3. The contour γ_R in Example 3

and note that the denominator of f vanishes precisely when $e^{\pi z} = -e^{-\pi z}$, that is, when $e^{2\pi z} = -1$. In other words, the only poles of f inside the rectangle are at the points $\alpha = i/2$ and $\beta = 3i/2$. To find the residue of f at α, we note that

$$\begin{aligned}(z-\alpha)f(z) &= e^{-2\pi iz\xi}\frac{2(z-\alpha)}{e^{\pi z}+e^{-\pi z}}\\ &= 2e^{-2\pi iz\xi}e^{\pi z}\frac{(z-\alpha)}{e^{2\pi z}-e^{2\pi\alpha}}.\end{aligned}$$

We recognize on the right the reciprocal of the difference quotient for the function $e^{2\pi z}$ at $z = \alpha$. Therefore

$$\lim_{z\to\alpha}(z-\alpha)f(z) = 2e^{-2\pi i\alpha\xi}e^{\pi\alpha}\frac{1}{2\pi e^{2\pi\alpha}} = \frac{e^{\pi\xi}}{\pi i},$$

which shows that f has a simple pole at α with residue $e^{\pi\xi}/(\pi i)$. Similarly, we find that f has a simple pole at β with residue $-e^{3\pi\xi}/(\pi i)$.

We dispense with the integrals of f on the vertical sides by showing that they go to zero as R tends to infinity. Indeed, if $z = R + iy$ with $0 \le y \le 2$, then

$$|e^{-2\pi iz\xi}| \le e^{4\pi|\xi|},$$

and

$$
\begin{aligned}
|\cosh \pi z| &= \left| \frac{e^{\pi z} + e^{-\pi z}}{2} \right| \\
&\geq \frac{1}{2} \left| |e^{\pi z}| - |e^{-\pi z}| \right| \\
&\geq \frac{1}{2}(e^{\pi R} - e^{-\pi R}) \\
&\to \infty \quad \text{ as } R \to \infty,
\end{aligned}
$$

which shows that the integral over the vertical segment on the right goes to 0 as $R \to \infty$. A similar argument shows that the integral of f over the vertical segment on the left also goes to 0 as $R \to \infty$. Finally, we see that if I denotes the integral we wish to calculate, then the integral of f over the top side of the rectangle (with the orientation from right to left) is simply $-e^{4\pi\xi}I$ where we have used the fact that $\cosh \pi\zeta$ is periodic with period $2i$. In the limit as R tends to infinity, the residue formula gives

$$
\begin{aligned}
I - e^{4\pi\xi} I &= 2\pi i \left(\frac{e^{\pi\xi}}{\pi i} - \frac{e^{3\pi\xi}}{\pi i} \right) \\
&= -2e^{2\pi\xi}(e^{\pi\xi} - e^{-\pi\xi}),
\end{aligned}
$$

and since $1 - e^{4\pi\xi} = -e^{2\pi\xi}(e^{2\pi\xi} - e^{-2\pi\xi})$, we find that

$$
I = 2\frac{e^{\pi\xi} - e^{-\pi\xi}}{e^{2\pi\xi} - e^{-2\pi\xi}} = 2\frac{e^{\pi\xi} - e^{-\pi\xi}}{(e^{\pi\xi} - e^{-\pi\xi})(e^{\pi\xi} + e^{-\pi\xi})} = \frac{2}{e^{\pi\xi} + e^{-\pi\xi}} = \frac{1}{\cosh \pi\xi}
$$

as claimed.

A similar argument actually establishes the following formula:

$$
\int_{-\infty}^{\infty} e^{-2\pi i x\xi} \frac{\sin \pi a}{\cosh \pi x + \cos \pi a}\, dx = \frac{2 \sinh 2\pi a\xi}{\sinh 2\pi\xi}
$$

whenever $0 < a < 1$, and where $\sinh z = (e^z - e^{-z})/2$. We have proved above the particular case $a = 1/2$. This identity can be used to determine an explicit formula for the Poisson kernel for the strip (see Problem 3 in Chapter 5 of Book I), or to prove the sum of two squares theorem, as we shall see in Chapter 10.

3 Singularities and meromorphic functions

Returning to Section 1, we see that we have described the analytical character of a function near a pole. We now turn our attention to the other types of isolated singularities.

Let f be a function holomorphic in an open set Ω except possibly at one point z_0 in Ω. If we can define f at z_0 in such a way that f becomes holomorphic in all of Ω, we say that z_0 is a **removable** singularity for f.

Theorem 3.1 (Riemann's theorem on removable singularities) *Suppose that f is holomorphic in an open set Ω except possibly at a point z_0 in Ω. If f is bounded on $\Omega - \{z_0\}$, then z_0 is a removable singularity.*

Proof. Since the problem is local we may consider a small disc D centered at z_0 and whose closure is contained in Ω. Let C denote the boundary circle of that disc with the usual positive orientation. We shall prove that if $z \in D$ and $z \neq z_0$, then under the assumptions of the theorem we have

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z} \, d\zeta. \tag{4}$$

Since an application of Theorem 5.4 in the previous chapter proves that the right-hand side of equation (4) defines a holomorphic function on *all* of D that agrees with $f(z)$ when $z \neq z_0$, this give us the desired extension.

To prove formula (4) we fix $z \in D$ with $z \neq z_0$ and use the familiar toy contour illustrated in Figure 4.

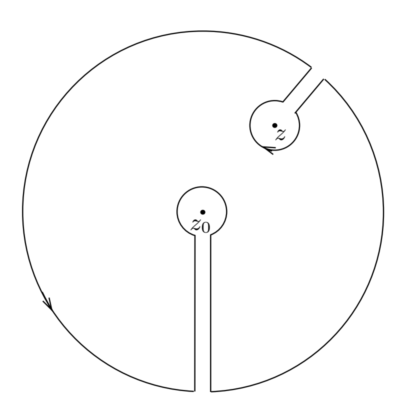


Figure 4. The multiple keyhole contour in the proof of Riemann's theorem

The multiple keyhole avoids the two points z and z_0. Letting the sides of the corridors get closer to each other, and finally overlap, in the limit

we get a cancellation:

$$\int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta + \int_{\gamma_\epsilon} \frac{f(\zeta)}{\zeta - z}\, d\zeta + \int_{\gamma'_\epsilon} \frac{f(\zeta)}{\zeta - z}\, d\zeta = 0\,,$$

where γ_ϵ and γ'_ϵ are small circles of radius ϵ with negative orientation and centered at z and z_0 respectively. Copying the argument used in the proof of the Cauchy integral formula in Section 4 of Chapter 2, we find that

$$\int_{\gamma_\epsilon} \frac{f(\zeta)}{\zeta - z}\, d\zeta = -2\pi i f(z).$$

For the second integral, we use the assumption that f is bounded and that since ϵ is small, ζ stays away from z, and therefore

$$\left| \int_{\gamma'_\epsilon} \frac{f(\zeta)}{\zeta - z}\, d\zeta \right| \leq C\epsilon.$$

Letting ϵ tend to 0 proves our contention and concludes the proof of the extension formula (4).

Surprisingly, we may deduce from Riemann's theorem a characterization of poles in terms of the behavior of the function in a neighborhood of a singularity.

Corollary 3.2 *Suppose that f has an isolated singularity at the point z_0. Then z_0 is a pole of f if and only if $|f(z)| \to \infty$ as $z \to z_0$.*

Proof. If z_0 is a pole, then we know that $1/f$ has a zero at z_0, and therefore $|f(z)| \to \infty$ as $z \to z_0$. Conversely, suppose that this condition holds. Then $1/f$ is bounded near z_0, and in fact $1/|f(z)| \to 0$ as $z \to z_0$. Therefore, $1/f$ has a removable singularity at z_0 and must vanish there. This proves the converse, namely that z_0 is a pole.

Isolated singularities belong to one of three categories:

- Removable singularities (f bounded near z_0)
- Pole singularities ($|f(z)| \to \infty$ as $z \to z_0$)
- Essential singularities.

By default, any singularity that is not removable or a pole is defined to be an **essential singularity**. For example, the function $e^{1/z}$ discussed at the very beginning of Section 1 has an essential singularity at

0. We already observed the wild behavior of this function near the origin. Contrary to the controlled behavior of a holomorphic function near a removable singularity or a pole, it is typical for a holomorphic function to behave erratically near an essential singularity. The next theorem clarifies this.

Theorem 3.3 (Casorati-Weierstrass) *Suppose f is holomorphic in the punctured disc $D_r(z_0) - \{z_0\}$ and has an essential singularity at z_0. Then, the image of $D_r(z_0) - \{z_0\}$ under f is dense in the complex plane.*

Proof. We argue by contradiction. Assume that the range of f is not dense, so that there exists $w \in \mathbb{C}$ and $\delta > 0$ such that

$$|f(z) - w| > \delta \quad \text{for all } z \in D_r(z_0) - \{z_0\}.$$

We may therefore define a new function on $D_r(z_0) - \{z_0\}$ by

$$g(z) = \frac{1}{f(z) - w},$$

which is holomorphic on the punctured disc and bounded by $1/\delta$. Hence g has a removable singularity at z_0 by Theorem 3.1. If $g(z_0) \neq 0$, then $f(z) - w$ is holomorphic at z_0, which contradicts the assumption that z_0 is an essential singularity. In the case that $g(z_0) = 0$, then $f(z) - w$ has a pole at z_0 also contradicting the nature of the singularity at z_0. The proof is complete.

In fact, Picard proved a much stronger result. He showed that under the hypothesis of the above theorem, the function f takes on every complex value infinitely many times with at most one exception. Although we shall not give a proof of this remarkable result, a simpler version of it will follow from our study of entire functions in a later chapter. See Exercise 11 in Chapter 5.

We now turn to functions with only isolated singularities that are poles. A function f on an open set Ω is **meromorphic** if there exists a sequence of points $\{z_0, z_1, z_2, \ldots\}$ that has no limit points in Ω, and such that

(i) the function f is holomorphic in $\Omega - \{z_0, z_1, z_2, \ldots\}$, and

(ii) f has poles at the points $\{z_0, z_1, z_2, \ldots\}$.

It is also useful to discuss functions that are meromorphic in the extended complex plane. If a function is holomorphic for all large values of

z, we can describe its behavior at infinity using the tripartite distinction we have used to classify singularities at finite values of z. Thus, if f is holomorphic for all large values of z, we consider $F(z) = f(1/z)$, which is now holomorphic in a deleted neighborhood of the origin. We say that f has a **pole at infinity** if F has a pole at the origin. Similarly, we can speak of f having an **essential singularity at infinity**, or a **removable singularity** (hence holomorphic) **at infinity** in terms of the corresponding behavior of F at 0. A meromorphic function in the complex plane that is either holomorphic at infinity or has a pole at infinity is said to be **meromorphic in the extended complex plane**.

At this stage we return to the principle mentioned at the beginning of the chapter. Here we can see it in its simplest form.

Theorem 3.4 *The meromorphic functions in the extended complex plane are the rational functions.*

Proof. Suppose that f is meromorphic in the extended plane. Then $f(1/z)$ has either a pole or a removable singularity at 0, and in either case it must be holomorphic in a deleted neighborhood of the origin, so that the function f can have only finitely many poles in the plane, say at $z_1, \ldots, z_n$. The idea is to subtract from f its principal parts at all its poles including the one at infinity. Near each pole $z_k \in \mathbb{C}$ we can write

$$f(z) = f_k(z) + g_k(z)\,,$$

where $f_k(z)$ is the principal part of f at z_k and g_k is holomorphic in a (full) neighborhood of z_k. In particular, f_k is a polynomial in $1/(z - z_k)$. Similarly, we can write

$$f(1/z) = \tilde{f}_\infty(z) + \tilde{g}_\infty(z)\,,$$

where $\tilde{g}_\infty$ is holomorphic in a neighborhood of the origin and $\tilde{f}_\infty$ is the principal part of $f(1/z)$ at 0, that is, a polynomial in $1/z$. Finally, let $f_\infty(z) = \tilde{f}_\infty(1/z)$.

We contend that the function $H = f - f_\infty - \sum_{k=1}^n f_k$ is entire and bounded. Indeed, near the pole z_k we subtracted the principal part of f so that the function H has a removable singularity there. Also, $H(1/z)$ is bounded for z near 0 since we subtracted the principal part of the pole at ∞. This proves our contention, and by Liouville's theorem we conclude that H is constant. From the definition of H, we find that f is a rational function, as was to be shown.

Note that as a consequence, a rational function is determined up to a multiplicative constant by prescribing the locations and multiplicities of its zeros and poles.

The Riemann sphere

The extended complex plane, which consists of $\mathbb{C}$ and the point at infinity, has a convenient geometric interpretation, which we briefly discuss here.

Consider the Euclidean space $\mathbb{R}^3$ with coordinates (X, Y, Z) where the XY-plane is identified with $\mathbb{C}$. We denote by $\mathbb{S}$ the sphere centered at $(0, 0, 1/2)$ and of radius $1/2$; this sphere is of unit diameter and lies on top of the origin of the complex plane as pictured in Figure 5. Also, we let $\mathcal{N} = (0, 0, 1)$ denote the north pole of the sphere.

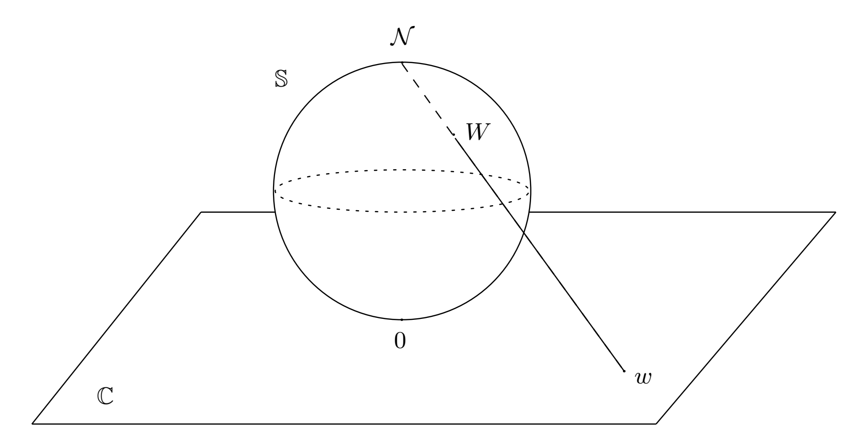


Figure 5. The Riemann sphere $\mathbb{S}$ and stereographic projection

Given any point $W = (X, Y, Z)$ on $\mathbb{S}$ different from the north pole, the line joining $\mathcal{N}$ and W intersects the XY-plane in a single point which we denote by $w = x + iy$; w is called the **stereographic projection** of W (see Figure 5). Conversely, given any point w in $\mathbb{C}$, the line joining $\mathcal{N}$ and $w = (x, y, 0)$ intersects the sphere at $\mathcal{N}$ and another point, which we call W. This geometric construction gives a bijective correspondence between points on the punctured sphere $\mathbb{S} - \{\mathcal{N}\}$ and the complex plane; it is described analytically by the formulas

$$x = \frac{X}{1 - Z} \quad \text{and} \quad y = \frac{Y}{1 - Z},$$

giving w in terms of W, and

$$X = \frac{x}{x^2 + y^2 + 1}, \quad Y = \frac{y}{x^2 + y^2 + 1}, \quad \text{and} \quad Z = \frac{x^2 + y^2}{x^2 + y^2 + 1}$$

giving W in terms of w. Intuitively, we have wrapped the complex plane onto the punctured sphere $\mathbb{S} - \{\mathcal{N}\}$.

As the point w goes to infinity in $\mathbb{C}$ (in the sense that $|w| \to \infty$) the corresponding point W on $\mathbb{S}$ comes arbitrarily close to $\mathcal{N}$. This simple observation makes $\mathcal{N}$ a natural candidate for the so-called "point at infinity." Identifying infinity with the point $\mathcal{N}$ on $\mathbb{S}$, we see that the extended complex plane can be visualized as the full two-dimensional sphere $\mathbb{S}$; this is the **Riemann sphere**. Since this construction takes the unbounded set $\mathbb{C}$ into the compact set $\mathbb{S}$ by adding one point, the Riemann sphere is sometimes called the **one-point compactification** of $\mathbb{C}$.

An important consequence of this interpretation is the following: although the point at infinity required special attention when considered separately from $\mathbb{C}$, it now finds itself on equal footing with all other points on $\mathbb{S}$. In particular, a meromorphic function on the extended complex plane can be thought of as a map from $\mathbb{S}$ to itself, where the image of a pole is now a tractable point on $\mathbb{S}$, namely $\mathcal{N}$. For these reasons (and others) the Riemann sphere provides good geometrical insight into the structure of $\mathbb{C}$ as well as the theory of meromorphic functions.

4 The argument principle and applications

We anticipate our discussion of the logarithm (in Section 6) with a few comments. In general, the function $\log f(z)$ is "multiple-valued" because it cannot be defined unambiguously on the set where $f(z) \neq 0$. However it is to be defined, it must equal $\log |f(z)| + i \arg f(z)$, where $\log |f(z)|$ is the usual real-variable logarithm of the positive quantity $|f(z)|$ (and hence is defined unambiguously), while $\arg f(z)$ is some determination of the argument (up to an additive integral multiple of 2π). Note that in any case, the derivative of $\log f(z)$ is $f'(z)/f(z)$ which is single-valued, and the integral

$$\int_\gamma \frac{f'(z)}{f(z)}\, dz$$

can be interpreted as the change in the argument of f as z traverses the curve γ. Moreover, assuming the curve is closed, this change of argument is determined entirely by the zeros and poles of f inside γ. We now formulate this fact as a precise theorem.

We begin with the observation that while the additivity formula

$$\log(f_1 f_2) = \log f_1 + \log f_2$$

fails in general (as we shall see below), the additivity can be restored to the corresponding derivatives. This is confirmed by the following

observation:

$$\frac{(f_1 f_2)'}{f_1 f_2} = \frac{f_1' f_2 + f_1 f_2'}{f_1 f_2} = \frac{f_1'}{f_1} + \frac{f_2'}{f_2},$$

which generalizes to

$$\frac{\left(\prod_{k=1}^N f_k\right)'}{\prod_{k=1}^N f_k} = \sum_{k=1}^N \frac{f_k'}{f_k}.$$

We apply this formula as follows. If f is holomorphic and has a zero of order n at z_0, we can write

$$f(z) = (z - z_0)^n g(z)\,,$$

where g is holomorphic and nowhere vanishing in a neighborhood of z_0, and therefore

$$\frac{f'(z)}{f(z)} = \frac{n}{z - z_0} + G(z)$$

where $G(z) = g'(z)/g(z)$. The conclusion is that if f has a zero of order n at z_0, then f'/f has a simple pole with residue n at z_0. Observe that a similar fact also holds if f has a pole of order n at z_0, that is, if $f(z) = (z - z_0)^{-n} h(z)$. Then

$$\frac{f'(z)}{f(z)} = \frac{-n}{z - z_0} + H(z).$$

Therefore, if f is meromorphic, the function f'/f will have simple poles at the zeros and poles of f, and the residue is simply the order of the zero of f or the negative of the order of the pole of f. As a result, an application of the residue formula gives the following theorem.

Theorem 4.1 (Argument principle) *Suppose f is meromorphic in an open set containing a circle C and its interior. If f has no poles and never vanishes on C, then*

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = \text{(number of zeros of } f \text{ inside } C\text{) minus}$$
$$\text{(number of poles of } f \text{ inside } C\text{)},$$

where the zeros and poles are counted with their multiplicities.

Corollary 4.2 *The above theorem holds for toy contours.*

As an application of the argument principle, we shall prove three theorems of interest in the general theory. The first, Rouché's theorem, is in some sense a continuity statement. It says that a holomorphic function can be perturbed slightly without changing the number of its zeros. Then, we prove the open mapping theorem, which states that holomorphic functions map open sets to open sets, an important property that again shows the special nature of holomorphic functions. Finally, the maximum modulus principle is reminiscent of (and in fact implies) the same property for harmonic functions: a non-constant holomorphic function on an open set Ω cannot attain its maximum in the interior of Ω.

Theorem 4.3 (Rouché's theorem) *Suppose that f and g are holomorphic in an open set containing a circle C and its interior. If*

$$|f(z)| > |g(z)| \quad \text{for all } z \in C,$$

then f and $f+g$ have the same number of zeros inside the circle C.

Proof. For $t \in [0,1]$ define

$$f_t(z) = f(z) + tg(z)$$

so that $f_0 = f$ and $f_1 = f+g$. Let n_t denote the number of zeros of f_t inside the circle counted with multiplicities, so that in particular, n_t is an integer. The condition $|f(z)| > |g(z)|$ for $z \in C$ clearly implies that f_t has no zeros on the circle, and the argument principle implies

$$n_t = \frac{1}{2\pi i}\int_C \frac{f_t'(z)}{f_t(z)}\, dz.$$

To prove that n_t is constant, it suffices to show that it is a continuous function of t. Then we could argue that if n_t were not constant, the intermediate value theorem would guarantee the existence of some $t_0 \in [0,1]$ with n_{t_0} not integral, contradicting the fact that $n_t \in \mathbb{Z}$ for all t.

To prove the continuity of n_t, we observe that $f_t'(z)/f_t(z)$ is jointly continuous for $t \in [0,1]$ and $z \in C$. This joint continuity follows from the fact that it holds for both the numerator and denominator, and our assumptions guarantee that $f_t(z)$ does not vanish on C. Hence n_t is integer-valued and continuous, and it must be constant. We conclude that $n_0 = n_1$, which is Rouché's theorem.

We now come to an important geometric property of holomorphic functions that arises when we consider them as mappings (that is, mapping regions in the complex plane to the complex plane).

A mapping is said to be **open** if it maps open sets to open sets.

Theorem 4.4 (Open mapping theorem) *If f is holomorphic and non-constant in a region Ω, then f is open.*

Proof. Let w_0 belong to the image of f, say $w_0 = f(z_0)$. We must prove that all points w near w_0 also belong to the image of f.

Define $g(z) = f(z) - w$ and write

$$\begin{aligned} g(z) &= (f(z) - w_0) + (w_0 - w) \\ &= F(z) + G(z). \end{aligned}$$

Now choose $\delta > 0$ such that the disc $|z - z_0| \leq \delta$ is contained in Ω and $f(z) \neq w_0$ on the circle $|z - z_0| = \delta$. We then select $\epsilon > 0$ so that we have $|f(z) - w_0| \geq \epsilon$ on the circle $|z - z_0| = \delta$. Now if $|w - w_0| < \epsilon$ we have $|F(z)| > |G(z)|$ on the circle $|z - z_0| = \delta$, and by Rouché's theorem we conclude that $g = F + G$ has a zero inside the circle since F has one.

The next result pertains to the size of a holomorphic function. We shall refer to the **maximum** of a holomorphic function f in an open set Ω as the maximum of its absolute value $|f|$ in Ω.

Theorem 4.5 (Maximum modulus principle) *If f is a non-constant holomorphic function in a region Ω, then f cannot attain a maximum in Ω.*

Proof. Suppose that f did attain a maximum at z_0. Since f is holomorphic it is an open mapping, and therefore, if $D \subset \Omega$ is a small disc centered at z_0, its image $f(D)$ is open and contains $f(z_0)$. This proves that there are points in $z \in D$ such that $|f(z)| > |f(z_0)|$, a contradiction.

Corollary 4.6 *Suppose that Ω is a region with compact closure $\overline{\Omega}$. If f is holomorphic on Ω and continuous on $\overline{\Omega}$ then*

$$\sup_{z \in \Omega} |f(z)| \leq \sup_{z \in \overline{\Omega} - \Omega} |f(z)|.$$

In fact, since $f(z)$ is continuous on the compact set $\overline{\Omega}$, then $|f(z)|$ attains its maximum in $\overline{\Omega}$; but this cannot be in Ω if f is non-constant. If f is constant, the conclusion is trivial.

Remark. The hypothesis that $\overline{\Omega}$ is compact (that is, bounded) is essential for the conclusion. We give an example related to considerations that we will take up in Chapter 4. Let Ω be the open first quadrant, bounded by the positive half-line $x \geq 0$ and the corresponding imaginary line $y \geq 0$. Consider $F(z) = e^{-iz^2}$. Then F is entire and clearly

continuous on $\overline{\Omega}$. Moreover $|F(z)| = 1$ on the two boundary lines $z = x$ and $z = iy$. However, $F(z)$ is unbounded in Ω, since for example, we have $F(z) = e^{r^2}$ if $z = r\sqrt{i} = re^{i\pi/4}$.

5 Homotopies and simply connected domains

The key to the general form of Cauchy's theorem, as well as the analysis of multiple-valued functions, is to understand in what regions we can define the primitive of a given holomorphic function. Note the relevance to the study of the logarithm, which arises as a primitive of $1/z$. The question is not just a local one, but is also global in nature. Its elucidation requires the notion of homotopy, and the resulting idea of simple-connectivity.

Let γ_0 and γ_1 be two curves in an open set Ω with common end-points. So if $\gamma_0(t)$ and $\gamma_1(t)$ are two parametrizations defined on $[a, b]$, we have

$$\gamma_0(a) = \gamma_1(a) = \alpha \quad \text{and} \quad \gamma_0(b) = \gamma_1(b) = \beta.$$

These two curves are said to be **homotopic in** Ω if for each $0 \leq s \leq 1$ there exists a curve $\gamma_s \subset \Omega$, parametrized by $\gamma_s(t)$ defined on $[a, b]$, such that for every s

$$\gamma_s(a) = \alpha \quad \text{and} \quad \gamma_s(b) = \beta,$$

and for all $t \in [a, b]$

$$\gamma_s(t)|_{s=0} = \gamma_0(t) \quad \text{and} \quad \gamma_s(t)|_{s=1} = \gamma_1(t).$$

Moreover, $\gamma_s(t)$ should be jointly continuous in $s \in [0, 1]$ and $t \in [a, b]$.

Loosely speaking, two curves are homotopic if one curve can be deformed into the other by a continuous transformation without ever leaving Ω (Figure 6).

Theorem 5.1 *If f is holomorphic in Ω, then*

$$\int_{\gamma_0} f(z)\, dz = \int_{\gamma_1} f(z)\, dz$$

whenever the two curves γ_0 and γ_1 are homotopic in Ω.

Proof. The key to the proof lies in showing that if two curves are close to each other and have the same end-points, then the integrals over these curves are equal. Recall that by definition, the function $F(s, t) = \gamma_s(t)$ is continuous on $[0, 1] \times [a, b]$. In particular, since the image of F, which we

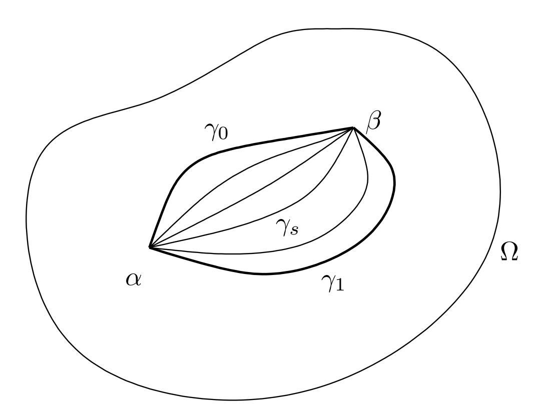


Figure 6. Homotopy of curves

denote by K, is compact, there exists $\epsilon > 0$ such that every disc of radius 3ϵ centered at a point in the image of F is completely contained in Ω. If not, for every $\ell \geq 0$, there exist points $z_\ell \in K$ and w_ℓ in the complement of Ω such that $|z_\ell - w_\ell| < 1/\ell$. By compactness of K, there exists a subsequence of $\{z_\ell\}$, say $\{z_{\ell_k}\}$, that converges to a point $z \in K \subset \Omega$. By construction, we must have $w_{\ell_k} \to z$ as well, and since $\{w_\ell\}$ lies in the complement of Ω which is closed, we must have $z \in \Omega^c$ as well. This is a contradiction.

Having found an ϵ with the desired property, we may, by the uniform continuity of F, select δ so that

$$\sup_{t \in [a,b]} |\gamma_{s_1}(t) - \gamma_{s_2}(t)| < \epsilon \qquad \text{whenever } |s_1 - s_2| < \delta.$$

Fix s_1 and s_2 with $|s_1 - s_2| < \delta$. We then choose discs $\{D_0, \ldots, D_n\}$ of radius 2ϵ, and consecutive points $\{z_0, \ldots, z_{n+1}\}$ on γ_{s_1} and $\{w_0, \ldots, w_{n+1}\}$ on γ_{s_2} such that the union of these discs covers both curves, and

$$z_i, z_{i+1}, w_i, w_{i+1} \in D_i.$$

The situation is illustrated in Figure 7.

Also, we choose $z_0 = w_0$ as the beginning end-point of the curves and $z_{n+1} = w_{n+1}$ as the common end-point. On each disc D_i, let F_i denote a primitive of f (Theorem 2.1, Chapter 2). On the intersection of D_i and D_{i+1}, F_i and F_{i+1} are two primitives of the same function, so they must differ by a constant, say c_i. Therefore

$$F_{i+1}(z_{i+1}) - F_i(z_{i+1}) = F_{i+1}(w_{i+1}) - F_i(w_{i+1}),$$

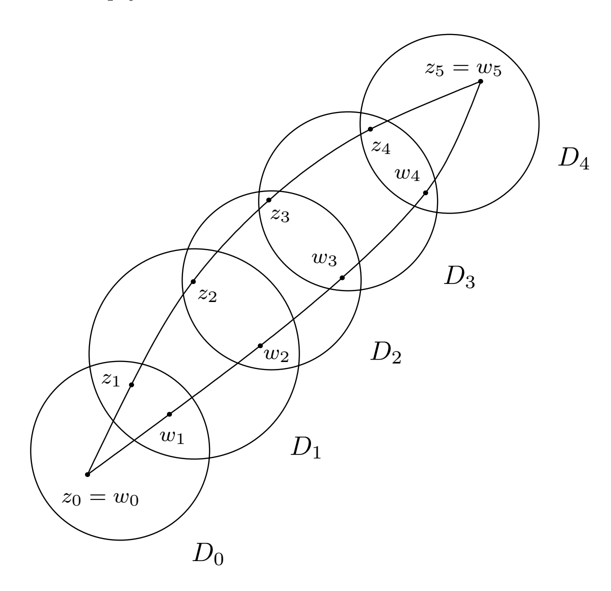


Figure 7. Covering two nearby curves with discs

hence

$$F_{i+1}(z_{i+1}) - F_{i+1}(w_{i+1}) = F_i(z_{i+1}) - F_i(w_{i+1}). \tag{5}$$

This implies

$$\begin{aligned}\int_{\gamma_{s_1}} f - \int_{\gamma_{s_2}} f &= \sum_{i=0}^{n}[F_i(z_{i+1}) - F_i(z_i)] - \sum_{i=0}^{n}[F_i(w_{i+1}) - F_i(w_i)] \\ &= \sum_{i=0}^{n} F_i(z_{i+1}) - F_i(w_{i+1}) - (F_i(z_i) - F_i(w_i)) \\ &= F_n(z_{n+1}) - F_n(w_{n+1}) - (F_0(z_0) - F_0(w_0)),\end{aligned}$$

because of the cancellations due to (5). Since γ_{s_1} and γ_{s_2} have the same beginning and end point, we have proved that

$$\int_{\gamma_{s_1}} f = \int_{\gamma_{s_2}} f.$$

By subdividing $[0,1]$ into subintervals $[s_i, s_{i+1}]$ of length less than δ, we may go from γ_0 to γ_1 by finitely many applications of the above argument, and the theorem is proved.

A region Ω in the complex plane is **simply connected** if any two pair of curves in Ω with the same end-points are homotopic.

EXAMPLE 1. A disc D is simply connected. In fact, if $\gamma_0(t)$ and $\gamma_1(t)$ are two curves lying in D, we can define $\gamma_s(t)$ by

$$\gamma_s(t) = (1-s)\gamma_0(t) + s\gamma_1(t).$$

Note that if $0 \leq s \leq 1$, then for each t, the point $\gamma_s(t)$ is on the segment joining $\gamma_0(t)$ and $\gamma_1(t)$, and so is in D. The same argument works if D is replaced by a rectangle, or more generally by any open convex set. (See Exercise 21.)

EXAMPLE 2. The slit plane $\Omega = \mathbb{C} - \{(-\infty, 0]\}$ is simply connected. For a pair of curves γ_0 and γ_1 in Ω, we write $\gamma_j(t) = r_j(t)e^{i\theta_j(t)}$ $(j = 0, 1)$ with $r_j(t)$ continuous and strictly positive, and $\theta_j(t)$ continuous with $|\theta_j(t)| < \pi$. Then, we can define $\gamma_s(t)$ as $r_s(t)e^{i\theta_s(t)}$ where

$$r_s(t) = (1-s)r_0(t) + sr_1(t) \quad \text{and} \quad \theta_s(t) = (1-s)\theta_0(t) + s\theta_1(t).$$

We then have $\gamma_s(t) \in \Omega$ whenever $0 \leq s \leq 1$.

EXAMPLE 3. With some effort one can show that the interior of a toy contour is simply connected. This requires that we divide the interior into several subregions. A general form of the argument is given in Exercise 4.

EXAMPLE 4. In contrast with the previous examples, the punctured plane $\mathbb{C} - \{0\}$ is not simply connected. Intuitively, consider two curves with the origin between them. It is impossible to continuously pass from one curve to the other without going over 0. A rigorous proof of this fact requires one further result, and will be given shortly.

Theorem 5.2 *Any holomorphic function in a simply connected domain has a primitive.*

Proof. Fix a point z_0 in Ω and define

$$F(z) = \int_\gamma f(w)\, dw$$

where the integral is taken over any curve in Ω joining z_0 to z. This definition is independent of the curve chosen, since Ω is simply connected,

and if $\tilde{\gamma}$ is another curve in Ω joining z_0 and z, we would have

$$\int_\gamma f(w)\,dw = \int_{\tilde{\gamma}} f(w)\,dw$$

by Theorem 5.1. Now we can write

$$F(z+h) - F(z) = \int_\eta f(w)\,dw$$

where η is the line segment joining z and $z+h$. Arguing as in the proof of Theorem 2.1 in Chapter 2, we find that

$$\lim_{h\to 0} \frac{F(z+h)-F(z)}{h} = f(z).$$

As a result, we obtain the following version of Cauchy's theorem.

Corollary 5.3 *If f is holomorphic in the simply connected region Ω, then*

$$\int_\gamma f(z)\,dz = 0$$

for any closed curve γ in Ω.

This is immediate from the existence of a primitive.

The fact that the punctured plane is not simply connected now follows rigorously from the observation that the integral of $1/z$ over the unit circle is $2\pi i$, and not 0.

6 The complex logarithm

Suppose we wish to define the logarithm of a non-zero complex number. If $z = re^{i\theta}$, and we want the logarithm to be the inverse to the exponential, then it is natural to set

$$\log z = \log r + i\theta.$$

Here and below, we use the convention that $\log r$ denotes the *standard*[1] logarithm of the positive number r. The trouble with the above definition is that θ is unique only up to an integer multiple of 2π. However,

[1]By the standard logarithm, we mean the natural logarithm of positive numbers that appears in elementary calculus.

for given z we can fix a choice of θ, and if z varies only a little, this determines the corresponding choice of θ uniquely (assuming we require that θ varies continuously with z). Thus "locally" we can give an unambiguous definition of the logarithm, but this will not work "globally." For example, if z starts at 1, and then winds around the origin and returns to 1, the logarithm does not return to its original value, but rather differs by an integer multiple of $2\pi i$, and therefore is not "single-valued." To make sense of the logarithm as a single-valued function, we must restrict the set on which we define it. This is the so-called choice of a **branch** or **sheet** of the logarithm.

Our discussion of simply connected domains given above leads to a natural global definition of a branch of the logarithm function.

Theorem 6.1 *Suppose that Ω is simply connected with $1 \in \Omega$, and $0 \notin \Omega$. Then in Ω there is a branch of the logarithm $F(z) = \log_\Omega(z)$ so that*

(i) *F is holomorphic in Ω,*

(ii) *$e^{F(z)} = z$ for all $z \in \Omega$,*

(iii) *$F(r) = \log r$ whenever r is a real number and near 1.*

In other words, each branch $\log_\Omega(z)$ is an extension of the standard logarithm defined for positive numbers.

Proof. We shall construct F as a primitive of the function $1/z$. Since $0 \notin \Omega$, the function $f(z) = 1/z$ is holomorphic in Ω. We define

$$\log_\Omega(z) = F(z) = \int_\gamma f(w)\, dw\,,$$

where γ is any curve in Ω connecting 1 to z. Since Ω is simply connected, this definition does not depend on the path chosen. Arguing as in the proof of Theorem 5.2, we find that F is holomorphic and $F'(z) = 1/z$ for all $z \in \Omega$. This proves (i). To prove (ii), it suffices to show that $ze^{-F(z)} = 1$. For that, we differentiate the left-hand side, obtaining

$$\frac{d}{dz}\left(ze^{-F(z)}\right) = e^{-F(z)} - zF'(z)e^{-F(z)} = (1 - zF'(z))e^{-F(z)} = 0.$$

Since Ω is connected we conclude, by Corollary 3.4 in Chapter 1, that $ze^{-F(z)}$ is constant. Evaluating this expression at $z = 1$, and noting that $F(1) = 0$, we find that this constant must be 1.

Finally, if r is real and close to 1 we can choose as a path from 1 to r a line segment on the real axis so that

$$F(r) = \int_1^r \frac{dx}{x} = \log r,$$

by the usual formula for the standard logarithm. This completes the proof of the theorem.

For example, in the slit plane $\Omega = \mathbb{C} - \{(-\infty, 0]\}$ we have the **principal branch** of the logarithm

$$\log z = \log r + i\theta$$

where $z = re^{i\theta}$ with $|\theta| < \pi$. (Here we drop the subscript Ω, and write simply $\log z$.) To prove this, we use the path of integration γ shown in Figure 8.

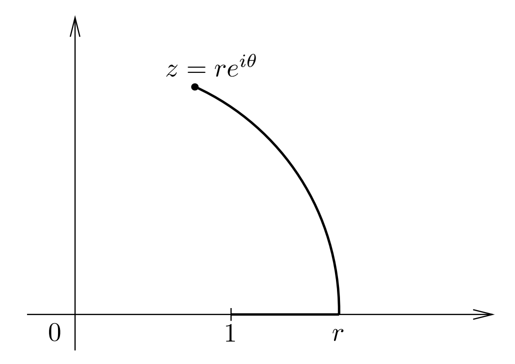


Figure 8. Path of integration for the principal branch of the logarithm

If $z = re^{i\theta}$ with $|\theta| < \pi$, the path consists of the line segment from 1 to r and the arc η from r to z. Then

$$\begin{aligned} \log z &= \int_1^r \frac{dx}{x} + \int_\eta \frac{dw}{w} \\ &= \log r + \int_0^\theta \frac{ire^{it}}{re^{it}}\, dt \\ &= \log r + i\theta. \end{aligned}$$

An important observation is that in general

$$\log(z_1 z_2) \neq \log z_1 + \log z_2.$$

For example, if $z_1 = e^{2\pi i/3} = z_2$, then for the principal branch of the logarithm, we have

$$\log z_1 = \log z_2 = \frac{2\pi i}{3},$$

and since $z_1 z_2 = e^{-2\pi i/3}$ we have

$$-\frac{2\pi i}{3} = \log(z_1 z_2) \neq \log z_1 + \log z_2.$$

Finally, for the principal branch of the logarithm the following Taylor expansion holds:

$$\log(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \cdots = -\sum_{n=1}^{\infty}(-1)^n \frac{z^n}{n} \quad \text{for } |z| < 1. \tag{6}$$

Indeed, the derivative of both sides equals $1/(1+z)$, so that they differ by a constant. Since both sides are equal to 0 at $z = 0$ this constant must be 0, and we have proved the desired Taylor expansion.

Having defined a logarithm on a simply connected domain, we can now define the powers z^α for any $\alpha \in \mathbb{C}$. If Ω is simply connected with $1 \in \Omega$ and $0 \notin \Omega$, we choose the branch of the logarithm with $\log 1 = 0$ as above, and define

$$z^\alpha = e^{\alpha \log z}.$$

Note that $1^\alpha = 1$, and that if $\alpha = 1/n$, then

$$(z^{1/n})^n = \prod_{k=1}^{n} e^{\frac{1}{n}\log z} = e^{\sum_{k=1}^{n} \frac{1}{n}\log z} = e^{\frac{n}{n}\log z} = e^{\log z} = z.$$

We know that every non-zero complex number w can be written as $w = e^z$. A generalization of this fact is given in the next theorem, which discusses the existence of $\log f(z)$ whenever f does not vanish.

Theorem 6.2 *If f is a nowhere vanishing holomorphic function in a simply connected region Ω, then there exists a holomorphic function g on Ω such that*

$$f(z) = e^{g(z)}.$$

The function $g(z)$ in the theorem can be denoted by $\log f(z)$, and determines a "branch" of that logarithm.

Proof. Fix a point z_0 in Ω, and define a function

$$g(z) = \int_\gamma \frac{f'(w)}{f(w)}\, dw + c_0,$$

where γ is any path in Ω connecting z_0 to z, and c_0 is a complex number so that $e^{c_0} = f(z_0)$. This definition is independent of the path γ since Ω is simply connected. Arguing as in the proof of Theorem 2.1, Chapter 2, we find that g is holomorphic with

$$g'(z) = \frac{f'(z)}{f(z)},$$

and a simple calculation gives

$$\frac{d}{dz}\left(f(z)e^{-g(z)}\right) = 0\,,$$

so that $f(z)e^{-g(z)}$ is constant. Evaluating this expression at z_0 we find $f(z_0)e^{-c_0} = 1$, so that $f(z) = e^{g(z)}$ for all $z \in \Omega$, and the proof is complete.

7 Fourier series and harmonic functions

In Chapter 4 we shall describe some interesting connections between complex function theory and Fourier analysis on the real line. The motivation for this study comes in part from the simple and direct relation between Fourier series on the circle and power series expansions of holomorphic functions in the disc, which we now investigate.

Suppose that f is holomorphic in a disc $D_R(z_0)$, so that f has a power series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

that converges in that disc.

Theorem 7.1 *The coefficients of the power series expansion of f are given by*

$$a_n = \frac{1}{2\pi r^n}\int_0^{2\pi} f(z_0 + re^{i\theta})e^{-in\theta}\,d\theta$$

for all $n \geq 0$ and $0 < r < R$. Moreover,

$$0 = \frac{1}{2\pi r^n}\int_0^{2\pi} f(z_0 + re^{i\theta})e^{-in\theta}\,d\theta$$

whenever $n < 0$.

Proof. Since $f^{(n)}(z_0) = a_n n!$, the Cauchy integral formula gives

$$a_n = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} \, d\zeta \,,$$

where γ is a circle of radius $0 < r < R$ centered at z_0 and with the positive orientation. Choosing $\zeta = z_0 + re^{i\theta}$ for the parametrization of this circle, we find that for $n \geq 0$

$$\begin{aligned} a_n &= \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{(z_0 + re^{i\theta} - z_0)^{n+1}} rie^{i\theta} \, d\theta \\ &= \frac{1}{2\pi r^n} \int_0^{2\pi} f(z_0 + re^{i\theta}) e^{-i(n+1)\theta} e^{i\theta} \, d\theta \\ &= \frac{1}{2\pi r^n} \int_0^{2\pi} f(z_0 + re^{i\theta}) e^{-in\theta} \, d\theta. \end{aligned}$$

Finally, even when $n < 0$, our calculation shows that we still have the identity

$$\frac{1}{2\pi r^n} \int_0^{2\pi} f(z_0 + re^{i\theta}) e^{-in\theta} \, d\theta = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} \, d\zeta.$$

Since $-n > 0$, the function $f(\zeta)(\zeta - z_0)^{-n-1}$ is holomorphic in the disc, and by Cauchy's theorem the last integral vanishes.

The interpretation of this theorem is as follows. Consider $f(z_0 + re^{i\theta})$ as the restriction to the circle of a holomorphic function on the closure of a disc centered at z_0 with radius r. Then its Fourier coefficients vanish if $n < 0$, while those for $n \geq 0$ are equal (up to a factor of r^n) to coefficients of the power series of the holomorphic function f. The property of the vanishing of the Fourier coefficients for $n < 0$ reveals another special characteristic of holomorphic functions (and in particular their restrictions to any circle).

Next, since $a_0 = f(z_0)$, we obtain the following corollary.

Corollary 7.2 (Mean-value property) *If f is holomorphic in a disc $D_R(z_0)$, then*

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) \, d\theta, \qquad \textit{for any } 0 < r < R.$$

Taking the real parts of both sides, we obtain the following consequence.

Corollary 7.3 *If f is holomorphic in a disc $D_R(z_0)$, and $u = \text{Re}(f)$, then*

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta, \qquad \textit{for any } 0 < r < R.$$

Recall that u is harmonic whenever f is holomorphic, and in fact, the above corollary is a property enjoyed by every harmonic function in the disc $D_R(z_0)$. This follows from Exercise 12 in Chapter 2, which shows that every harmonic function in a disc is the real part of a holomorphic function in that disc.

8 Exercises

1. Using Euler's formula

$$\sin \pi z = \frac{e^{i\pi z} - e^{-i\pi z}}{2i},$$

show that the complex zeros of $\sin \pi z$ are exactly at the integers, and that they are each of order 1.

Calculate the residue of $1/\sin \pi z$ at $z = n \in \mathbb{Z}$.

2. Evaluate the integral

$$\int_{-\infty}^{\infty} \frac{dx}{1 + x^4}.$$

Where are the poles of $1/(1 + z^4)$?

3. Show that

$$\int_{-\infty}^{\infty} \frac{\cos x}{x^2 + a^2}\, dx = \pi \frac{e^{-a}}{a}, \qquad \text{for } a > 0.$$

4. Show that

$$\int_{-\infty}^{\infty} \frac{x \sin x}{x^2 + a^2}\, dx = \pi e^{-a}, \qquad \text{for all } a > 0.$$

5. Use contour integration to show that

$$\int_{-\infty}^{\infty} \frac{e^{-2\pi i x \xi}}{(1 + x^2)^2}\, dx = \frac{\pi}{2}(1 + 2\pi|\xi|)e^{-2\pi|\xi|}$$

for all ξ real.

6. Show that

$$\int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^{n+1}} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \cdot \pi.$$

7. Prove that

$$\int_0^{2\pi} \frac{d\theta}{(a+\cos\theta)^2} = \frac{2\pi a}{(a^2-1)^{3/2}}, \qquad \text{whenever } a > 1.$$

8. Prove that

$$\int_0^{2\pi} \frac{d\theta}{a + b\cos\theta} = \frac{2\pi}{\sqrt{a^2-b^2}}$$

if $a > |b|$ and $a, b \in \mathbb{R}$.

9. Show that

$$\int_0^1 \log(\sin \pi x)\, dx = -\log 2.$$

[Hint: Use the contour shown in Figure 9.]

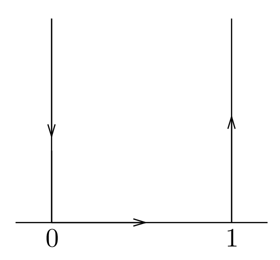


Figure 9. Contour in Exercise 9

10. Show that if $a > 0$, then

$$\int_0^{\infty} \frac{\log x}{x^2+a^2}\, dx = \frac{\pi}{2a} \log a.$$

[Hint: Use the contour in Figure 10.]

11. Show that if $|a| < 1$, then

$$\int_0^{2\pi} \log|1 - ae^{i\theta}|\, d\theta = 0.$$

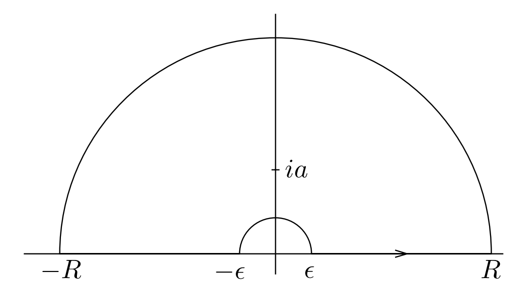


Figure 10. Contour in Exercise 10

Then, prove that the above result remains true if we assume only that $|a| \leq 1$.

12. Suppose u is not an integer. Prove that

$$\sum_{n=-\infty}^{\infty} \frac{1}{(u+n)^2} = \frac{\pi^2}{(\sin \pi u)^2}$$

by integrating

$$f(z) = \frac{\pi \cot \pi z}{(u+z)^2}$$

over the circle $|z| = R_N = N + 1/2$ (N integral, $N \geq |u|$), adding the residues of f inside the circle, and letting N tend to infinity.

Note. Two other derivations of this identity, using Fourier series, were given in Book I.

13. Suppose $f(z)$ is holomorphic in a punctured disc $D_r(z_0) - \{z_0\}$. Suppose also that

$$|f(z)| \leq A|z - z_0|^{-1+\epsilon}$$

for some $\epsilon > 0$, and all z near z_0. Show that the singularity of f at z_0 is removable.

14. Prove that all entire functions that are also injective take the form $f(z) = az + b$ with $a, b \in \mathbb{C}$, and $a \neq 0$.

[Hint: Apply the Casorati-Weierstrass theorem to $f(1/z)$.]

15. Use the Cauchy inequalities or the maximum modulus principle to solve the following problems:

(a) Prove that if f is an entire function that satisfies

$$\sup_{|z|=R} |f(z)| \leq AR^k + B$$

for all $R > 0$, and for some integer $k \geq 0$ and some constants $A, B > 0$, then f is a polynomial of degree $\leq k$.

(b) Show that if f is holomorphic in the unit disc, is bounded, and converges uniformly to zero in the sector $\theta < \arg z < \varphi$ as $|z| \to 1$, then $f = 0$.

(c) Let $w_1, \ldots, w_n$ be points on the unit circle in the complex plane. Prove that there exists a point z on the unit circle such that the product of the distances from z to the points w_j, $1 \leq j \leq n$, is at least 1. Conclude that there exists a point w on the unit circle such that the product of the distances from w to the points w_j, $1 \leq j \leq n$, is exactly equal to 1.

(d) Show that if the real part of an entire function f is bounded, then f is constant.

16. Suppose f and g are holomorphic in a region containing the disc $|z| \leq 1$. Suppose that f has a simple zero at $z = 0$ and vanishes nowhere else in $|z| \leq 1$. Let

$$f_\epsilon(z) = f(z) + \epsilon g(z).$$

Show that if ϵ is sufficiently small, then

(a) $f_\epsilon(z)$ has a unique zero in $|z| \leq 1$, and

(b) if z_ϵ is this zero, the mapping $\epsilon \mapsto z_\epsilon$ is continuous.

17. Let f be non-constant and holomorphic in an open set containing the closed unit disc.

(a) Show that if $|f(z)| = 1$ whenever $|z| = 1$, then the image of f contains the unit disc. [Hint: One must show that $f(z) = w_0$ has a root for every $w_0 \in \mathbb{D}$. To do this, it suffices to show that $f(z) = 0$ has a root (why?). Use the maximum modulus principle to conclude.]

(b) If $|f(z)| \geq 1$ whenever $|z| = 1$ and there exists a point $z_0 \in \mathbb{D}$ such that $|f(z_0)| < 1$, then the image of f contains the unit disc.

18. Give another proof of the Cauchy integral formula

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

using homotopy of curves.

[Hint: Deform the circle C to a small circle centered at z, and note that the quotient $(f(\zeta) - f(z))/(\zeta - z)$ is bounded.]

19. Prove the maximum principle for harmonic functions, that is:

(a) If u is a non-constant real-valued harmonic function in a region Ω, then u cannot attain a maximum (or a minimum) in Ω.

(b) Suppose that Ω is a region with compact closure $\overline{\Omega}$. If u is harmonic in Ω and continuous in $\overline{\Omega}$, then

$$\sup_{z\in\Omega}|u(z)| \leq \sup_{z\in\overline{\Omega}-\Omega}|u(z)|.$$

[Hint: To prove the first part, assume that u attains a local maximum at z_0. Let f be holomorphic near z_0 with $u = \text{Re}(f)$, and show that f is not open. The second part follows directly from the first.]

20. This exercise shows how the mean square convergence dominates the uniform convergence of analytic functions. If U is an open subset of $\mathbb{C}$ we use the notation

$$\|f\|_{L^2(U)} = \left(\int_U |f(z)|^2\, dxdy\right)^{1/2}$$

for the mean square norm, and

$$\|f\|_{L^\infty(U)} = \sup_{z\in U}|f(z)|$$

for the sup norm.

(a) If f is holomorphic in a neighborhood of the disc $D_r(z_0)$, show that for any $0 < s < r$ there exists a constant $C > 0$ (which depends on s and r) such that

$$\|f\|_{L^\infty(D_s(z_0))} \leq C\|f\|_{L^2(D_r(z_0))}\,.$$

(b) Prove that if $\{f_n\}$ is a Cauchy sequence of holomorphic functions in the mean square norm $\|\cdot\|_{L^2(U)}$, then the sequence $\{f_n\}$ converges uniformly on every compact subset of U to a holomorphic function.

[Hint: Use the mean-value property.]

21. Certain sets have geometric properties that guarantee they are simply connected.

(a) An open set $\Omega \subset \mathbb{C}$ is **convex** if for any two points in Ω, the straight line segment between them is contained in Ω. Prove that a convex open set is simply connected.

(b) More generally, an open set $\Omega \subset \mathbb{C}$ is **star-shaped** if there exists a point $z_0 \in \Omega$ such that for any $z \in \Omega$, the straight line segment between z and z_0 is contained in Ω. Prove that a star-shaped open set is simply connected. Conclude that the slit plane $\mathbb{C} - \{(-\infty, 0]\}$ (and more generally any sector, convex or not) is simply connected.

(c) What are other examples of open sets that are simply connected?

22. Show that there is no holomorphic function f in the unit disc $\mathbb{D}$ that extends continuously to $\partial\mathbb{D}$ such that $f(z) = 1/z$ for $z \in \partial\mathbb{D}$.

9 Problems

1.* Consider a holomorphic map on the unit disc $f : \mathbb{D} \to \mathbb{C}$ which satisfies $f(0) = 0$. By the open mapping theorem, the image $f(\mathbb{D})$ contains a small disc centered at the origin. We then ask: does there exist $r > 0$ such that for *all* $f : \mathbb{D} \to \mathbb{C}$ with $f(0) = 0$, we have $D_r(0) \subset f(\mathbb{D})$?

(a) Show that with no further restrictions on f, no such r exists. It suffices to find a sequence of functions $\{f_n\}$ holomorphic in $\mathbb{D}$ such that $1/n \notin f(\mathbb{D})$. Compute $f_n'(0)$, and discuss.

(b) Assume in addition that f also satisfies $f'(0) = 1$. Show that despite this new assumption, there exists no $r > 0$ satisfying the desired condition.

[Hint: Try $f_\epsilon(z) = \epsilon(e^{z/\epsilon} - 1)$.]

The Koebe-Bieberbach theorem states that if in addition to $f(0) = 0$ and $f'(0) = 1$ we also assume that f is injective, then such an r exists and the best possible value is $r = 1/4$.

(c) As a first step, show that if $h(z) = \frac{1}{z} + c_0 + c_1 z + c_2 z^2 + \cdots$ is analytic and injective for $0 < |z| < 1$, then $\sum_{n=1}^{\infty} n|c_n|^2 \leq 1$.

[Hint: Calculate the area of the complement of $h(D_\rho(0) - \{0\})$ where $0 < \rho < 1$, and let $\rho \to 1$.]

(d) If $f(z) = z + a_2 z^2 + \cdots$ satisfies the hypotheses of the theorem, show that there exists another function g satisfying the hypotheses of the theorem such that $g^2(z) = f(z^2)$.

[Hint: $f(z)/z$ is nowhere vanishing so there exists ψ such that $\psi^2(z) = f(z)/z$ and $\psi(0) = 1$. Check that $g(z) = z\psi(z^2)$ is injective.]

(e) With the notation of the previous part, show that $|a_2| \leq 2$, and that equality holds if and only if

$$f(z) = \frac{z}{(1 - e^{i\theta} z)^2} \qquad \text{for some } \theta \in \mathbb{R}.$$

[Hint: What is the power series expansion of $1/g(z)$? Use part (c).]

(f) If $h(z) = \frac{1}{z} + c_0 + c_1 z + c_2 z^2 + \cdots$ is injective on $\mathbb{D}$ and avoids the values z_1 and z_2, show that $|z_1 - z_2| \leq 4$.

[Hint: Look at the second coefficient in the power series expansion of $1/(h(z) - z_j)$.]

(g) Complete the proof of the theorem. [Hint: If f avoids w, then $1/f$ avoids 0 and $1/w$.]

2. Let u be a harmonic function in the unit disc that is continuous on its closure. Deduce Poisson's integral formula

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-|z_0|^2}{|e^{i\theta}-z_0|^2} u(e^{i\theta})\, d\theta \qquad \text{for } |z_0|<1$$

from the special case $z_0 = 0$ (the mean value theorem). Show that if $z_0 = re^{i\varphi}$, then

$$\frac{1-|z_0|^2}{|e^{i\theta}-z_0|^2} = \frac{1-r^2}{1-2r\cos(\theta-\varphi)+r^2} = P_r(\theta-\varphi),$$

and we recover the expression for the Poisson kernel derived in the exercises of the previous chapter.

[Hint: Set $u_0(z) = u(T(z))$ where

$$T(z) = \frac{z_0 - z}{1-\overline{z_0}z}.$$

Prove that u_0 is harmonic. Then apply the mean value theorem to u_0, and make a change of variables in the integral.]

3. If $f(z)$ is holomorphic in the deleted neighborhood $\{0<|z-z_0|<r\}$ and has a pole of order k at z_0, then we can write

$$f(z) = \frac{a_{-k}}{(z-z_0)^k} + \cdots + \frac{a_{-1}}{(z-z_0)} + g(z)$$

where g is holomorphic in the disc $\{|z-z_0|<r\}$. There is a generalization of this expansion that holds even if z_0 is an essential singularity. This is a special case of the **Laurent series expansion**, which is valid in an even more general setting.

Let f be holomorphic in a region containing the annulus $\{z : r_1 \le |z-z_0| \le r_2\}$ where $0<r_1<r_2$. Then,

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z-z_0)^n$$

where the series converges absolutely in the interior of the annulus. To prove this, it suffices to write

$$f(z) = \frac{1}{2\pi i}\int_{C_{r_2}} \frac{f(\zeta)}{\zeta - z}\, d\zeta - \frac{1}{2\pi i}\int_{C_{r_1}} \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

when $r_1 < |z-z_0| < r_2$, and argue as in the proof of Theorem 4.4, Chapter 2. Here C_{r_1} and C_{r_2} are the circles bounding the annulus.

4.* Suppose Ω is a bounded region. Let L be a (two-way infinite) line that intersects Ω. Assume that $\Omega \cap L$ is an interval I. Choosing an orientation for L, we can define Ω_l and Ω_r to be the subregions of Ω lying strictly to the left or right of L, with $\Omega = \Omega_l \cup I \cup \Omega_r$ a disjoint union. If Ω_l and Ω_r are simply connected, then Ω is simply connected.

5.* Let

$$g(z) = \frac{1}{2\pi i} \int_{-M}^{M} \frac{h(x)}{x - z}\, dx$$

where h is continuous and supported in $[-M, M]$.

(a) Prove that the function g is holomorphic in $\mathbb{C} - [-M, M]$, and vanishes at infinity, that is, $\lim_{|z|\to\infty} |g(z)| = 0$. Moreover, the "jump" of g across $[-M, M]$ is h, that is,

$$h(x) = \lim_{\epsilon \to 0, \epsilon > 0} g(x + i\epsilon) - g(x - i\epsilon).$$

[Hint: Express the difference $g(x + i\epsilon) - g(x - i\epsilon)$ in terms of a convolution of h with the Poisson kernel.]

(b) If h satisfies a mild smoothness condition, for instance a Hölder condition with exponent α, that is, $|h(x) - h(y)| \leq C|x - y|^{\alpha}$ for some $C > 0$ and all $x, y \in [-M, M]$, then $g(x + i\epsilon)$ and $g(x - i\epsilon)$ converge uniformly to functions $g_+(x)$ and $g_-(x)$ as $\epsilon \to 0$. Then, g can be characterized as the unique holomorphic function that satisfies:

(i) g is holomorphic outside $[-M, M]$,

(ii) g vanishes at infinity,

(iii) $g(x + i\epsilon)$ and $g(x - i\epsilon)$ converge uniformly as $\epsilon \to 0$ to functions $g_+(x)$ and $g_-(x)$ with

$$g_+(x) - g_-(x) = h(x).$$

[Hint: If G is another function satisfying these conditions, $g - G$ is entire.]

4 The Fourier Transform

> Raymond Edward Alan Christopher Paley, Fellow of Trinity College, Cambridge, and International Research Fellow at the Massachusetts Institute of Technology and at Harvard University, was killed by an avalanche on April 7, 1933, while skiing in the vicinity of Banff, Alberta. Although only twenty-six years of age, he was already recognized as the ablest of the group of young English mathematicians who have been inspired by the genius of G. H. Hardy and J. E. Littlewood. In a group notable for its brilliant technique, no one had developed this technique to a higher degree than Paley. Nevertheless he should not be thought of primarily as a technician, for with this ability he combined creative power of the first order. As he himself was wont to say, technique without "rugger tactics" will not get one far, and these rugger tactics he practiced to a degree that was characteristic of his forthright and vigorous nature.
>
> Possessed of an extraordinary capacity for making friends and for scientific collaboration, Paley believed that the inspiration of continual interchange of ideas stimulates each collaborator to accomplish more than he would alone. Only the exceptional man works well with a partner, but Paley had collaborated successfully with many, including Littlewood, Pólya, Zygmund, and Wiener.
>
> *N. Wiener,* 1933

If f is a function on $\mathbb{R}$ that satisfies appropriate regularity and decay conditions, then its Fourier transform is defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} \, dx, \qquad \xi \in \mathbb{R}$$

and its counterpart, the Fourier inversion formula, holds

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi, \qquad x \in \mathbb{R}.$$

The Fourier transform (including its d-dimensional variants), plays a ba-

sic role in analysis, as the reader of Book I is aware. Here we want to illustrate the intimate and fruitful connection between the one-dimensional theory of the Fourier transform and complex analysis. The main theme (stated somewhat imprecisely) is as follows: for a function f initially defined on the real line, the possibility of extending it to a holomorphic function is closely related to the very rapid (for example, exponential) decay at infinity of its Fourier transform $\hat{f}$. We elaborate on this theme in two stages.

First, we assume that f can be analytically continued in a horizontal strip containing the real axis, and has "moderate decrease" at infinity,[1] so that the integral defining the Fourier transform $\hat{f}$ converges. As a result, we conclude that $\hat{f}$ decreases exponentially at infinity; it also follows directly that the Fourier inversion formula holds. Moreover one can easily obtain from these considerations the Poisson summation formula $\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$. Incidentally, all these theorems are elegant consequences of contour integration.

At a second stage, we take as our starting point the validity of the Fourier inversion formula, which holds if we assume that both f and $\hat{f}$ are of moderate decrease, without making any assumptions on the analyticity of f. We then ask a simple but natural question: What are the conditions on f so that its Fourier transform is supported in a bounded interval, say $[-M, M]$? This is a basic problem that, as one notices, can be stated without any reference to notions of complex analysis. However, it can be resolved only in terms of the holomorphic properties of the function f. The condition, given by the Paley-Wiener theorem, is that there be a holomorphic extension of f to $\mathbb{C}$ that satisfies the growth condition

$$|f(z)| \leq Ae^{2\pi M|z|} \quad \text{for some constant } A > 0.$$

Functions satisfying this condition are said to be of **exponential type**.

Observe that the condition that $\hat{f}$ vanishes outside a compact set can be viewed as an extreme version of a decay property at infinity, and so the above theorem clearly falls within the context of the theme indicated above.

In all these matters a decisive technique will consist in shifting the contour of integration, that is the real line, within the boundaries of a horizontal strip. This will take advantage of the special behavior of $e^{-2\pi iz\xi}$ when z has a non-zero imaginary part. Indeed, when z is real this exponential remains bounded and oscillates, while if $\text{Im}(z) \neq 0$, it will

[1]We say that a function f is of **moderate decrease** if f is continuous and there exists $A > 0$ so that $|f(x)| \leq A/(1 + x^2)$ for all $x \in \mathbb{R}$. A more restrictive condition is that $f \in \mathcal{S}$, the Schwartz space of testing functions, which also implies that $\hat{f}$ belongs to $\mathcal{S}$. See Book I for more details.

have exponential decay or exponential increase, depending on whether the product $\xi \text{Im}(z)$ is negative or positive.

1 The class $\mathfrak{F}$

The weakest decay condition imposed on functions in our study of the Fourier transform in Book I was that of moderate decrease. There, we proved the Fourier inversion and Poisson summation formulas under the hypothesis that f and $\hat{f}$ satisfy

$$|f(x)| \leq \frac{A}{1+x^2} \quad \text{and} \quad |\hat{f}(\xi)| \leq \frac{A'}{1+\xi^2}$$

for some positive constants A, A' and all $x, \xi \in \mathbb{R}$. We were led to consider this class of functions because of various examples such as the Poisson kernel

$$P_y(x) = \frac{1}{\pi} \frac{y}{y^2 + x^2}$$

for $y > 0$, which played a fundamental role in the solution of the Dirichlet problem for the steady-state heat equation in the upper half-plane. There we had $\widehat{P_y}(\xi) = e^{-2\pi y |\xi|}$.

In the present context, we introduce a class of functions particularly suited to the goal we have set: proving theorems about the Fourier transform using complex analysis. Moreover, this class will be large enough to contain many of the important applications we have in mind.

For each $a > 0$ we denote by $\mathfrak{F}_a$ the class of all functions f that satisfy the following two conditions:

(i) The function f is holomorphic in the horizontal strip

$$S_a = \{z \in \mathbb{C} : \ |\text{Im}(z)| < a\}.$$

(ii) There exists a constant $A > 0$ such that

$$|f(x+iy)| \leq \frac{A}{1+x^2} \quad \text{for all } x \in \mathbb{R} \text{ and } |y| < a.$$

In other words, $\mathfrak{F}_a$ consists of those holomorphic functions on S_a that are of moderate decay on each horizontal line $\text{Im}(z) = y$, uniformly in $-a < y < a$. For example, $f(z) = e^{-\pi z^2}$ belongs to $\mathfrak{F}_a$ for all a. Also, the function

$$f(z) = \frac{1}{\pi} \frac{c}{c^2 + z^2},$$

which has simple poles at $z = \pm ci$, belongs to $\mathfrak{F}_a$ for all $0 < a < c$.

Another example is provided by $f(z) = 1/\cosh \pi z$, which belongs to $\mathfrak{F}_a$ whenever $|a| < 1/2$. This function, as well as one of its fundamental properties, was already discussed in Example 3, Section 2.1 of Chapter 3.

Note also that a simple application of the Cauchy integral formulas shows that if $f \in \mathfrak{F}_a$, then for every n, the nth derivative of f belongs to $\mathfrak{F}_b$ for all b with $0 < b < a$ (Exercise 2).

Finally, we denote by $\mathfrak{F}$ the class of all functions that belong to $\mathfrak{F}_a$ for some a.

Remark. The condition of moderate decrease can be weakened somewhat by replacing the order of decrease of $A/(1+x^2)$ by $A/(1+|x|^{1+\epsilon})$ for any $\epsilon > 0$. As the reader will observe, many of the results below remain unchanged with this less restrictive condition.

2 Action of the Fourier transform on $\mathfrak{F}$

Here we prove three theorems, including the Fourier inversion and Poisson summation formulas, for functions in $\mathfrak{F}$. The idea behind all three proofs is the same: contour integration. Thus the approach used will be different from that of the corresponding results in Book I.

Theorem 2.1 *If f belongs to the class $\mathfrak{F}_a$ for some $a > 0$, then $|\hat{f}(\xi)| \leq Be^{-2\pi b|\xi|}$ for any $0 \leq b < a$.*

Proof. Recall that $\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\xi}\,dx$. The case $b = 0$ simply says that $\hat{f}$ is bounded, which follows at once from the integral defining $\hat{f}$, the assumption that f is of moderate decrease, and the fact that the exponential is bounded by 1.

Now suppose $0 < b < a$ and assume first that $\xi > 0$. The main step consists of shifting the contour of integration, that is the real line, down by b. More precisely, consider the contour in Figure 1 as well as the function $g(z) = f(z)e^{-2\pi i z\xi}$.

We claim that as R tends to infinity, the integrals of g over the two vertical sides converge to zero. For example, the integral over the vertical segment on the left can be estimated by

$$\begin{aligned}
\left| \int_{-R-ib}^{-R} g(z)\,dz \right| &\leq \int_0^b |f(-R-it)e^{-2\pi i(-R-it)\xi}|\,dt \\
&\leq \int_0^b \frac{A}{R^2} e^{-2\pi t\xi}\,dt \\
&= O(1/R^2).
\end{aligned}$$

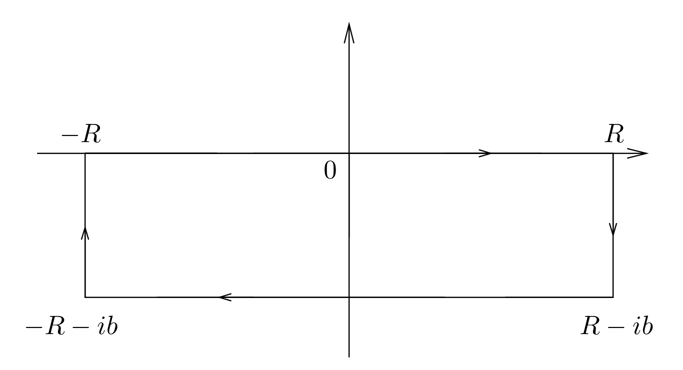


Figure 1. The contour in the proof of Theorem 2.1 when $\xi > 0$

A similar estimate for the other side proves our claim. Therefore, by Cauchy's theorem applied to the large rectangle, we find in the limit as R tends to infinity that

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x - ib)e^{-2\pi i(x-ib)\xi}\, dx, \tag{1}$$

which leads to the estimate

$$|\hat{f}(\xi)| \leq \int_{-\infty}^{\infty} \frac{A}{1+x^2} e^{-2\pi b\xi}\, dx \leq Be^{-2\pi b\xi},$$

where B is a suitable constant. A similar argument for $\xi < 0$, but this time shifting the real line up by b, allows us to finish the proof of the theorem.

This result says that whenever $f \in \mathfrak{F}$, then $\hat{f}$ has rapid decay at infinity. We remark that the further we can extend f (that is, the larger a), then the larger we can choose b, hence the better the decay. We shall return to this circle of ideas in Section 3, where we describe those f for which $\hat{f}$ has the ultimate decay condition: compact support.

Since $\hat{f}$ decreases rapidly on $\mathbb{R}$, the integral in the Fourier inversion formula makes sense, and we now turn to the complex analytic proof of this identity.

Theorem 2.2 *If $f \in \mathfrak{F}$, then the Fourier inversion holds, namely*

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi i x\xi}\, d\xi \quad \textit{for all } x \in \mathbb{R}.$$

Besides contour integration, the proof of the theorem requires a simple identity, which we isolate.

Lemma 2.3 *If A is positive and B is real, then* $\int_0^\infty e^{-(A+iB)\xi}\,d\xi = \frac{1}{A+iB}$.

Proof. Since $A > 0$ and $B \in \mathbb{R}$, we have $|e^{-(A+iB)\xi}| = e^{-A\xi}$, and the integral converges. By definition

$$\int_0^\infty e^{-(A+iB)\xi}\,d\xi = \lim_{R\to\infty}\int_0^R e^{-(A+iB)\xi}\,d\xi.$$

However,

$$\int_0^R e^{-(A+iB)\xi}\,d\xi = \left[-\frac{e^{-(A+iB)\xi}}{A+iB}\right]_0^R,$$

which tends to $1/(A+iB)$ as R tends to infinity.

We can now prove the inversion theorem. Once again, the sign of ξ matters, so we begin by writing

$$\int_{-\infty}^\infty \hat{f}(\xi)e^{2\pi i x\xi}\,d\xi = \int_{-\infty}^0 \hat{f}(\xi)e^{2\pi i x\xi}\,d\xi + \int_0^\infty \hat{f}(\xi)e^{2\pi i x\xi}\,d\xi.$$

For the second integral we argue as follows. Say $f \in \mathfrak{F}_a$ and choose $0 < b < a$. Arguing as the proof of Theorem 2.1, or simply using equation (1), we get

$$\hat{f}(\xi) = \int_{-\infty}^\infty f(u-ib)e^{-2\pi i(u-ib)\xi}\,du,$$

so that with an application of the lemma and the convergence of the integration in ξ, we find

$$\begin{aligned}
\int_0^\infty \hat{f}(\xi)e^{2\pi i x\xi}\,d\xi &= \int_0^\infty\int_{-\infty}^\infty f(u-ib)e^{-2\pi i(u-ib)\xi}e^{2\pi i x\xi}\,du\,d\xi \\
&= \int_{-\infty}^\infty f(u-ib)\int_0^\infty e^{-2\pi i(u-ib-x)\xi}\,d\xi\,du \\
&= \int_{-\infty}^\infty f(u-ib)\frac{1}{2\pi b+2\pi i(u-x)}\,du \\
&= \frac{1}{2\pi i}\int_{-\infty}^\infty \frac{f(u-ib)}{u-ib-x}\,du \\
&= \frac{1}{2\pi i}\int_{L_1}\frac{f(\zeta)}{\zeta-x}\,d\zeta,
\end{aligned}$$

where L_1 denotes the line $\{u - ib : u \in \mathbb{R}\}$ traversed from left to right. (In other words, L_1 is the real line shifted down by b.) For the integral when $\xi < 0$, a similar calculation gives

$$\int_{-\infty}^{0} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi = -\frac{1}{2\pi i} \int_{L_2} \frac{f(\zeta)}{\zeta - x} \, d\zeta,$$

where L_2 is the real line shifted up by b, with orientation from left to right. Now given $x \in \mathbb{R}$, consider the contour γ_R in Figure 2.

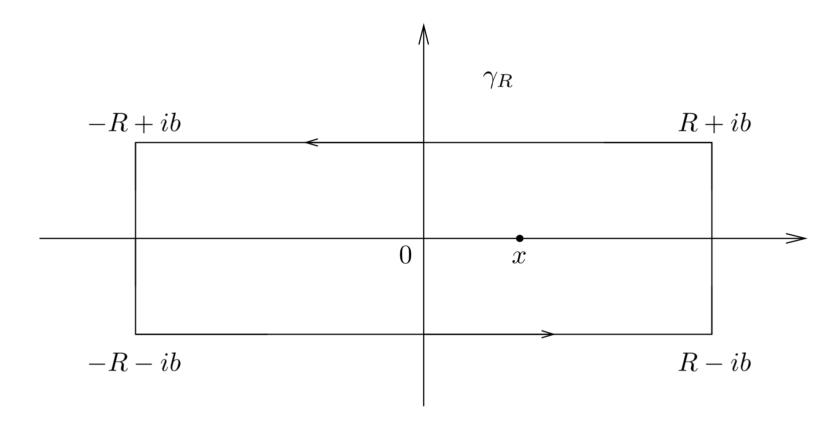


Figure 2. The contour γ_R in the proof of Theorem 2.2

The function $f(\zeta)/(\zeta - x)$ has a simple pole at x with residue $f(x)$, so the residue formula gives

$$f(x) = \frac{1}{2\pi i} \int_{\gamma_R} \frac{f(\zeta)}{\zeta - x} \, d\zeta.$$

Letting R tend to infinity, one checks easily that the integral over the vertical sides goes to 0 and therefore, combining with the previous results, we get

$$\begin{aligned} f(x) &= \frac{1}{2\pi i} \int_{L_1} \frac{f(\zeta)}{\zeta - x} \, d\zeta - \frac{1}{2\pi i} \int_{L_2} \frac{f(\zeta)}{\zeta - x} \, d\zeta \\ &= \int_0^{\infty} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi + \int_{-\infty}^{0} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi \\ &= \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi, \end{aligned}$$

and the theorem is proved.

The last of our three theorems is the Poisson summation formula.

Theorem 2.4 *If $f \in \mathfrak{F}$, then*

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n).$$

Proof. Say $f \in \mathfrak{F}_a$ and choose some b satisfying $0 < b < a$. The function $1/(e^{2\pi iz} - 1)$ has simple poles with residue $1/(2\pi i)$ at the integers. Thus $f(z)/(e^{2\pi iz} - 1)$ has simple poles at the integers n, with residues $f(n)/2\pi i$. We may therefore apply the residue formula to the contour γ_N in Figure 3 where N is an integer.

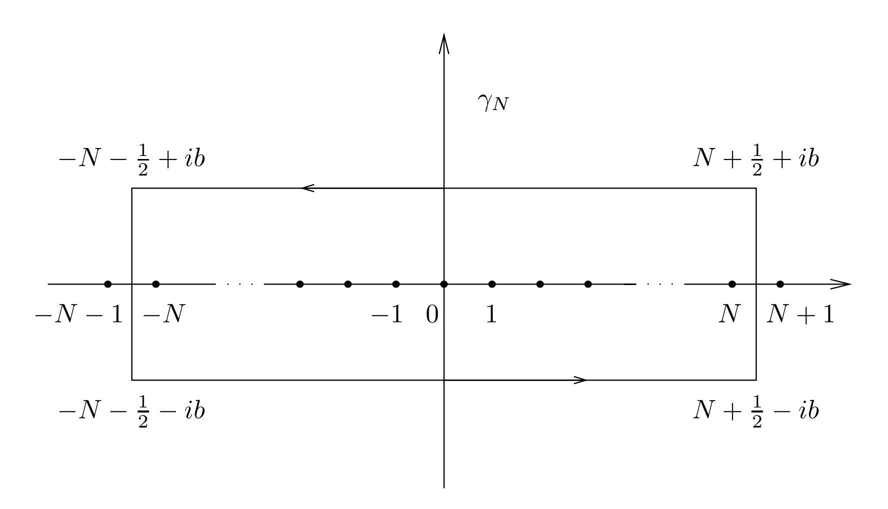


Figure 3. The contour γ_N in the proof of Theorem 2.4

This yields

$$\sum_{|n|\leq N} f(n) = \int_{\gamma_N} \frac{f(z)}{e^{2\pi iz} - 1}\, dz.$$

Letting N tend to infinity, and recalling that f has moderate decrease, we see that the sum converges to $\sum_{n\in\mathbb{Z}} f(n)$, and also that the integral over the vertical segments goes to 0. Therefore, in the limit we get

$$\sum_{n\in\mathbb{Z}} f(n) = \int_{L_1} \frac{f(z)}{e^{2\pi iz} - 1}\, dz - \int_{L_2} \frac{f(z)}{e^{2\pi iz} - 1}\, dz \tag{2}$$

where L_1 and L_2 are the real line shifted down and up by b, respectively.

Now we use the fact that if $|w| > 1$, then

$$\frac{1}{w-1} = w^{-1} \sum_{n=0}^{\infty} w^{-n}$$

to see that on L_1 (where $|e^{2\pi iz}| > 1$) we have

$$\frac{1}{e^{2\pi iz}-1} = e^{-2\pi iz} \sum_{n=0}^{\infty} e^{-2\pi inz}.$$

Also if $|w| < 1$, then

$$\frac{1}{w-1} = -\sum_{n=0}^{\infty} w^{n}$$

so that on L_2

$$\frac{1}{e^{2\pi iz}-1} = -\sum_{n=0}^{\infty} e^{2\pi inz}.$$

Substituting these observations in (2) we find that

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} f(n) &= \int_{L_1} f(z) \left(e^{-2\pi iz} \sum_{n=0}^{\infty} e^{-2\pi inz} \right) dz + \int_{L_2} f(z) \left(\sum_{n=0}^{\infty} e^{2\pi inz} \right) dz \\
&= \sum_{n=0}^{\infty} \int_{L_1} f(z) e^{-2\pi i(n+1)z}\, dz + \sum_{n=0}^{\infty} \int_{L_2} f(z) e^{2\pi inz}\, dz \\
&= \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-2\pi i(n+1)x}\, dx + \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} f(x) e^{2\pi inx}\, dz \\
&= \sum_{n=0}^{\infty} \hat{f}(n+1) + \sum_{n=0}^{\infty} \hat{f}(-n) \\
&= \sum_{n\in\mathbb{Z}} \hat{f}(n),
\end{aligned}$$

where we have shifted L_1 and L_2 back to the real line according to equation (1) and its analogue for the shift down.

The Poisson summation formula has many far-reaching consequences, and we close this section by deriving several interesting identities that are of importance for later applications.

First, we recall the calculation in Example 1, Chapter 2, which showed that the function $e^{-\pi x^2}$ was its own Fourier transform:

$$\int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx = e^{-\pi\xi^2}.$$

For fixed values of $t > 0$ and $a \in \mathbb{R}$, the change of variables $x \mapsto t^{1/2}(x+a)$ in the above integral shows that the Fourier transform of the function $f(x) = e^{-\pi t(x+a)^2}$ is $\hat{f}(\xi) = t^{-1/2} e^{-\pi\xi^2/t} e^{2\pi i a\xi}$. Applying the Poisson summation formula to the pair f and $\hat{f}$ (which belong to $\mathfrak{F}$) provides the following relation:

$$\sum_{n=-\infty}^{\infty} e^{-\pi t(n+a)^2} = \sum_{n=-\infty}^{\infty} t^{-1/2} e^{-\pi n^2/t} e^{2\pi i n a}. \tag{3}$$

This identity has noteworthy consequences. For instance, the special case $a = 0$ is the transformation law for a version of the "theta function": if we define ϑ for $t > 0$ by the series $\vartheta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t}$, then the relation (3) says precisely that

$$\vartheta(t) = t^{-1/2}\vartheta(1/t) \quad \text{for } t > 0. \tag{4}$$

This equation will be used in Chapter 6 to derive the key functional equation of the Riemann zeta function, and this leads to its analytic continuation. Also, the general case $a \in \mathbb{R}$ will be used in Chapter 10 to establish a corresponding law for the more general Jacobi theta function Θ.

For another application of the Poisson summation formula we recall that we proved in Example 3, Chapter 3, that the function $1/\cosh \pi x$ was also its own Fourier transform:

$$\int_{-\infty}^{\infty} \frac{e^{-2\pi i x\xi}}{\cosh \pi x}\, dx = \frac{1}{\cosh \pi\xi}.$$

This implies that if $t > 0$ and $a \in \mathbb{R}$, then the Fourier transform of the function $f(x) = e^{-2\pi i a x}/\cosh(\pi x/t)$ is $\hat{f}(\xi) = t/\cosh(\pi(\xi + a)t)$, and the Poisson summation formula yields

$$\sum_{n=-\infty}^{\infty} \frac{e^{-2\pi i a n}}{\cosh(\pi n/t)} = \sum_{n=-\infty}^{\infty} \frac{t}{\cosh(\pi(n+a)t)}. \tag{5}$$

This formula will be used in Chapter 10 in the context of the two-squares theorem.

3 Paley-Wiener theorem

In this section we change our point of view somewhat: we do not suppose any analyticity of f, but we do assume the validity of the Fourier inversion formula

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi i x\xi}\, d\xi \quad \text{if} \quad \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\xi}\, dx,$$

under the conditions $|f(x)| \leq A/(1+x^2)$ and $|\hat{f}(\xi)| \leq A'/(1+\xi^2)$. For a proof of the inversion formula under these conditions, we refer the reader to Chapter 5 in Book I.

We start by pointing out a partial converse to Theorem 2.1.

Theorem 3.1 *Suppose $\hat{f}$ satisfies the decay condition $|\hat{f}(\xi)| \leq Ae^{-2\pi a|\xi|}$ for some constants $a, A > 0$. Then $f(x)$ is the restriction to $\mathbb{R}$ of a function $f(z)$ holomorphic in the strip $S_b = \{z \in \mathbb{C} : \ |\text{Im}(z)| < b\}$, for any $0 < b < a$.*

Proof. Define

$$f_n(z) = \int_{-n}^{n} \hat{f}(\xi)e^{2\pi i \xi z}\, d\xi,$$

and note that f_n is entire by Theorem 5.4 in Chapter 2. Observe also that $f(z)$ may be defined for all z in the strip S_b by

$$f(z) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi i \xi z}\, d\xi\,,$$

because the integral converges absolutely by our assumption on $\hat{f}$: it is majorized by

$$A\int_{-\infty}^{\infty} e^{-2\pi a|\xi|}e^{2\pi b|\xi|}\, d\xi\,,$$

which is finite if $b < a$. Moreover, for $z \in S_b$

$$\begin{aligned} |f(z) - f_n(z)| &\leq A\int_{|\xi|\geq n} e^{-2\pi a|\xi|}e^{2\pi b|\xi|}\, d\xi \\ &\to 0 \quad \text{as } n \to \infty, \end{aligned}$$

and thus the sequence $\{f_n\}$ converges to f uniformly in S_b, which, by Theorem 5.2 in Chapter 2, proves the theorem.

We digress briefly to make the following observation.

Corollary 3.2 *If* $\hat{f}(\xi) = O(e^{-2\pi a|\xi|})$ *for some* $a > 0$*, and* f *vanishes in a non-empty open interval, then* $f = 0$*.*

Since by the theorem f is analytic in a region containing the real line, the corollary is a consequence of Theorem 4.8 in Chapter 2. In particular, we recover the fact proved in Exercise 21, Chapter 5 in Book I, namely that f and $\hat{f}$ cannot both have compact support unless $f = 0$.

The Paley-Wiener theorem goes a step further than the previous theorem, and describes the nature of those functions whose Fourier transforms are supported in a given interval $[-M, M]$.

Theorem 3.3 *Suppose* f *is continuous and of moderate decrease on* $\mathbb{R}$*. Then,* f *has an extension to the complex plane that is entire with* $|f(z)| \leq Ae^{2\pi M|z|}$ *for some* $A > 0$*, if and only if* $\hat{f}$ *is supported in the interval* $[-M, M]$*.*

One direction is simple. Suppose $\hat{f}$ is supported in $[-M, M]$. Then both f and $\hat{f}$ have moderate decrease, and the Fourier inversion formula applies

$$f(x) = \int_{-M}^{M} \hat{f}(\xi) e^{2\pi i \xi x}\, d\xi.$$

Since the range of integration is finite, we may replace x by the complex variable z in the integral, thereby defining a complex-valued function on $\mathbb{C}$ by

$$g(z) = \int_{-M}^{M} \hat{f}(\xi) e^{2\pi i \xi z}\, d\xi.$$

By construction $g(z) = f(z)$ if z is real, and g is holomorphic by Theorem 5.4 in Chapter 2. Finally, if $z = x + iy$, we have

$$\begin{aligned} |g(z)| &\leq \int_{-M}^{M} |\hat{f}(\xi)| e^{-2\pi \xi y}\, d\xi \\ &\leq Ae^{2\pi M|z|}. \end{aligned}$$

The converse result requires a little more work. It starts with the observation that if $\hat{f}$ were supported in $[-M, M]$, then the argument above would give the stronger bound $|f(z)| \leq Ae^{2\pi|y|}$ instead of what we assume, that is $|f(z)| \leq Ae^{2\pi|z|}$. The idea is then to try to reduce to the better situation, where this stronger bound holds. However, this is not quite enough because we need in addition a (moderate) decay as $x \to \infty$

(when $y \neq 0$) to deal with the convergence of certain integrals at infinity. Thus we begin by also assuming this further property of f, and then we remove the additional assumptions, one step at a time.

Step 1. We first assume that f is holomorphic in the complex plane, and satisfies the following condition regarding decay in x and growth in y:

$$|f(x+iy)| \leq A' \frac{e^{2\pi M|y|}}{1+x^2}. \tag{6}$$

We then prove under this stronger assumption that $\hat{f}(\xi) = 0$ if $|\xi| > M$. To see this, we first suppose that $\xi > M$ and write

$$\begin{aligned}\hat{f}(\xi) &= \int_{-\infty}^{\infty} f(x)e^{-2\pi i\xi x}\,dx \\ &= \int_{-\infty}^{\infty} f(x-iy)e^{-2\pi i\xi(x-iy)}\,dx.\end{aligned}$$

Here we have shifted the real line down by an amount $y > 0$ using the standard argument (equation (1)). Putting absolute values gives the bound

$$\begin{aligned}|\hat{f}(\xi)| &\leq A' \int_{-\infty}^{\infty} \frac{e^{2\pi My - 2\pi\xi y}}{1+x^2}\,dx \\ &\leq Ce^{-2\pi y(\xi - M)}.\end{aligned}$$

Letting y tend to infinity, and recalling that $\xi - M > 0$, proves that $\hat{f}(\xi) = 0$. A similar argument, shifting the contour up by $y > 0$, proves that $\hat{f}(\xi) = 0$ whenever $\xi < -M$.

Step 2. We relax condition (6) by assuming only that f satisfies

$$|f(x+iy)| \leq Ae^{2\pi M|y|}. \tag{7}$$

This is still a stronger condition than in the theorem, but it is weaker than (6). Suppose first that $\xi > M$, and for $\epsilon > 0$ consider the following auxiliary function

$$f_\epsilon(z) = \frac{f(z)}{(1+i\epsilon z)^2}.$$

We observe that the quantity $1/(1+i\epsilon z)^2$ has absolute value less than or equal to 1 in the closed lower half-plane (including the real line) and

converges to 1 as ϵ tends to 0. In particular, this shows that $\widehat{f_\epsilon}(\xi) \to \hat{f}(\xi)$ as $\epsilon \to 0$ since we may write

$$|\widehat{f_\epsilon}(\xi) - \hat{f}(\xi)| \leq \int_{-\infty}^{\infty} |f(x)| \left[\frac{1}{(1+i\epsilon x)^2} - 1\right] dx,$$

and recall that f has moderate decrease on $\mathbb{R}$.

But for each fixed ϵ, we have

$$|f_\epsilon(x+iy)| \leq A'' \frac{e^{2\pi M|y|}}{1+x^2},$$

so by Step 1 we must have $\widehat{f_\epsilon}(\xi) = 0$, and hence $\hat{f}(\xi) = 0$ after passing to the limit as $\epsilon \to 0$. A similar argument applies if $\xi < -M$, although we must now argue in the upper half-plane, and use the factor $1/(1 - i\epsilon z)^2$ instead.

Step 3. To conclude the proof, it suffices to show that the conditions in the theorem imply condition (7) in Step 2. In fact, after dividing by an appropriate constant, it suffices to show that if $|f(x)| \leq 1$ for all real x, and $|f(z)| \leq e^{2\pi M|z|}$ for all complex z, then

$$|f(x+iy)| \leq e^{2\pi M|y|}.$$

This will follow from an ingenious and very useful idea of Phragmén and Lindelöf that allows one to adapt the maximum modulus principle to various unbounded regions. The particular result we need is as follows.

Theorem 3.4 *Suppose F is a holomorphic function in the sector*

$$S = \{z : -\pi/4 < \arg z < \pi/4\}$$

that is continuous on the closure of S. Assume $|F(z)| \leq 1$ on the boundary of the sector, and that there are constants $C, c > 0$ such that $|F(z)| \leq Ce^{c|z|}$ for all z in the sector. Then

$$|F(z)| \leq 1 \quad \textit{for all } z \in S.$$

In other words, if F is bounded by 1 on the boundary of S and has no more than a reasonable amount of growth, then F is actually bounded everywhere by 1. That some restriction on the growth of F is necessary follows from a simple observation. Consider the function $F(z) = e^{z^2}$. Then F is bounded by 1 on the boundary of S, but if x is real, $F(x)$ is unbounded as $x \to \infty$. We now give the proof of Theorem 3.4.

Proof. The idea is to subdue the "enemy" function e^{z^2} and turn it to our advantage: in brief, one modifies e^{z^2} by replacing it by e^{z^α} with $\alpha < 2$. For simplicity we take the case $\alpha = 3/2$.

If $\epsilon > 0$, let

$$F_\epsilon(z) = F(z)e^{-\epsilon z^{3/2}}.$$

Here we have chosen the principal branch of the logarithm to define $z^{3/2}$ so that if $z = re^{i\theta}$ (with $-\pi < \theta < \pi$), then $z^{3/2} = r^{3/2}e^{3i\theta/2}$. Hence F_ϵ is holomorphic in S and continuous up to its boundary. Also

$$|e^{-\epsilon z^{3/2}}| = e^{-\epsilon r^{3/2}\cos(3\theta/2)};$$

and since $-\pi/4 < \theta < \pi/4$ in the sector, we get the inequalities

$$-\frac{\pi}{2} < -\frac{3\pi}{8} < \frac{3\theta}{2} < \frac{3\pi}{8} < \frac{\pi}{2},$$

and therefore $\cos(3\theta/2)$ is strictly positive in the sector. This, together with the fact that $|F(z)| \leq Ce^{c|z|}$, shows that $F_\epsilon(z)$ decreases rapidly in the closed sector as $|z| \to \infty$, and in particular F_ϵ is bounded. We claim that in fact $|F_\epsilon(z)| \leq 1$ for all $z \in \overline{S}$, where $\overline{S}$ denotes the closure of S. To prove this, we define

$$M = \sup_{z\in\overline{S}} |F_\epsilon(z)|.$$

Assuming F is not identically zero, let $\{w_j\}$ be a sequence of points such that $|F_\epsilon(w_j)| \to M$. Since $M \neq 0$ and F_ϵ decays to 0 as $|z|$ becomes large in the sector, w_j cannot escape to infinity, and we conclude that this sequence accumulates to a point $w \in \overline{S}$. By the maximum principle, w cannot be an interior point of S, so w lies on its boundary. But on the boundary, we have first $|F(z)| \leq 1$ by assumption, and also $|e^{-\epsilon z^{3/2}}| \leq 1$, so that $M \leq 1$, and the claim is proved.

Finally, we may let ϵ tend to 0 to conclude the proof of the theorem.

Further generalizations of the Phragmén-Lindelöf theorem are included in Exercise 9 and Problem 3.

We must now use this result to conclude the proof of the Paley-Wiener theorem, that is, show that if $|f(x)| \leq 1$ and $|f(z)| \leq e^{2\pi M|z|}$, then $|f(z)| \leq e^{2\pi M|y|}$. First, note that the sector in the Phragmén-Lindelöf theorem can be rotated, say to the first quadrant $Q = \{z = x + iy :\ x > 0,\ y > 0\}$, and the result remains the same. Then, we consider

$$F(z) = f(z)e^{2\pi iMz},$$

and note that F is bounded by 1 on the positive real and positive imaginary axes. Since we also have $|F(z)| \leq Ce^{c|z|}$ in the quadrant, we conclude by the Phragmén-Lindelöf theorem that $|F(z)| \leq 1$ for all z in Q, which implies $|f(z)| \leq e^{2\pi My}$. A similar argument for the other quadrants concludes Step 3 as well as the proof of the Paley-Wiener theorem.

We conclude with another version of the idea behind the Paley-Wiener theorem, this time characterizing the functions whose Fourier transform vanishes for all negative ξ.

Theorem 3.5 *Suppose f and $\hat{f}$ have moderate decrease. Then $\hat{f}(\xi) = 0$ for all $\xi < 0$ if and only if f can be extended to a continuous and bounded function in the closed upper half-plane $\{z = x + iy : y \geq 0\}$ with f holomorphic in the interior.*

Proof. First assume $\hat{f}(\xi) = 0$ for $\xi < 0$. By the Fourier inversion formula

$$f(x) = \int_0^\infty \hat{f}(\xi) e^{2\pi i x\xi}\, d\xi\,,$$

and we can extend f when $z = x + iy$ with $y \geq 0$ by

$$f(z) = \int_0^\infty \hat{f}(\xi) e^{2\pi i z\xi}\, d\xi\,.$$

Notice that the integral converges and that

$$|f(z)| \leq A \int_0^\infty \frac{d\xi}{1 + \xi^2} < \infty\,,$$

which proves the boundedness of f. The uniform convergence of

$$f_n(z) = \int_0^n \hat{f}(\xi) e^{2\pi i x\xi}\, d\xi$$

to $f(z)$ in the closed half-plane establishes the continuity of f there, and its holomorphicity in the interior.

For the converse, we argue in the spirit of the proof of Theorem 3.3. For ϵ and δ positive, we set

$$f_{\epsilon,\delta}(z) = \frac{f(z + i\delta)}{(1 - i\epsilon z)^2}.$$

Then $f_{\epsilon,\delta}$ is holomorphic in a region containing the closed upper half-plane. One also shows as before, using Cauchy's theorem, that $\widehat{f_{\epsilon,\delta}}(\xi) = 0$

for all $\xi < 0$. Then, passing to the limit successively, one has $\widehat{f_{\epsilon,0}}(\xi) = 0$ for $\xi < 0$, and finally $\hat{f}(\xi) = \widehat{f_{0,0}}(\xi) = 0$ for all $\xi < 0$.

Remark. The reader should note a certain analogy between the above theorem and Theorem 7.1 in Chapter 3. Here we are dealing with a function holomorphic in the upper half-plane, and there with a function holomorphic in a disc. In the present case the Fourier transform vanishes when $\xi < 0$, and in the earlier case, the Fourier coefficients vanish when $n < 0$.

4 Exercises

1. Suppose f is continuous and of moderate decrease, and $\hat{f}(\xi) = 0$ for all $\xi \in \mathbb{R}$. Show that $f = 0$ by completing the following outline:

(a) For each fixed real number t consider the two functions

$$A(z) = \int_{-\infty}^{t} f(x)e^{-2\pi iz(x-t)}\,dx \quad \text{and} \quad B(z) = -\int_{t}^{\infty} f(x)e^{-2\pi iz(x-t)}\,dx.$$

Show that $A(\xi) = B(\xi)$ for all $\xi \in \mathbb{R}$.

(b) Prove that the function F equal to A in the closed upper half-plane, and B in the lower half-plane, is entire and bounded, thus constant. In fact, show that $F = 0$.

(c) Deduce that

$$\int_{-\infty}^{t} f(x)\,dx = 0,$$

for all t, and conclude that $f = 0$.

2. If $f \in \mathfrak{F}_a$ with $a > 0$, then for any positive integer n one has $f^{(n)} \in \mathfrak{F}_b$ whenever $0 \le b < a$.

[Hint: Modify the solution to Exercise 8 in Chapter 2.]

3. Show, by contour integration, that if $a > 0$ and $\xi \in \mathbb{R}$ then

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{a}{a^2 + x^2} e^{-2\pi ix\xi}\,dx = e^{-2\pi a|\xi|},$$

and check that

$$\int_{-\infty}^{\infty} e^{-2\pi a|\xi|}e^{2\pi i\xi x}\,d\xi = \frac{1}{\pi}\frac{a}{a^2 + x^2}.$$

4. Suppose Q is a polynomial of degree ≥ 2 with distinct roots, none lying on the real axis. Calculate

$$\int_{-\infty}^{\infty} \frac{e^{-2\pi i x \xi}}{Q(x)}\, dx, \qquad \xi \in \mathbb{R}$$

in terms of the roots of Q. What happens when several roots coincide?

[Hint: Consider separately the cases $\xi < 0$, $\xi = 0$, and $\xi > 0$. Use residues.]

5. More generally, let $R(x) = P(x)/Q(x)$ be a rational function with (degree Q) $\geq$ (degree$P$)+2 and $Q(x) \neq 0$ on the real axis.

(a) Prove that if $\alpha_1, \ldots, \alpha_k$ are the roots of R in the upper half-plane, then there exists polynomials $P_j(\xi)$ of degree less than the multiplicity of α_j so that

$$\int_{-\infty}^{\infty} R(x) e^{-2\pi i x \xi}\, dx = \sum_{j=1}^{k} P_j(\xi) e^{-2\pi i \alpha_j \xi}, \qquad \text{when } \xi < 0.$$

(b) In particular, if $Q(z)$ has no zeros in the upper half-plane, then $\int_{-\infty}^{\infty} R(x) e^{-2\pi i x \xi}\, dx = 0$ for $\xi < 0$.

(c) Show that similar results hold in the case $\xi > 0$.

(d) Show that

$$\int_{-\infty}^{\infty} R(x) e^{-2\pi i x \xi}\, dx = O(e^{-a|\xi|}), \qquad \xi \in \mathbb{R}$$

as $|\xi| \to \infty$ for some $a > 0$. Determine the best possible a's in terms of the roots of R.

[Hint: For part (a), use residues. The powers of ξ appear when one differentiates the function $f(z) = R(z)e^{-2\pi i z \xi}$ (as in the formula of Theorem 1.4 in the previous chapter). For part (c) argue in the lower half-plane.]

6. Prove that

$$\frac{1}{\pi} \sum_{n=-\infty}^{\infty} \frac{a}{a^2 + n^2} = \sum_{n=-\infty}^{\infty} e^{-2\pi a |n|}$$

whenever $a > 0$. Hence show that the sum equals $\coth \pi a$.

7. The Poisson summation formula applied to specific examples often provides interesting identities.

(a) Let τ be fixed with $\mathrm{Im}(\tau) > 0$. Apply the Poisson summation formula to

$$f(z) = (\tau + z)^{-k},$$

where k is an integer ≥ 2, to obtain

$$\sum_{n=-\infty}^{\infty} \frac{1}{(\tau+n)^k} = \frac{(-2\pi i)^k}{(k-1)!} \sum_{m=1}^{\infty} m^{k-1} e^{2\pi i m \tau}.$$

(b) Set $k = 2$ in the above formula to show that if $\mathrm{Im}(\tau) > 0$, then

$$\sum_{n=-\infty}^{\infty} \frac{1}{(\tau+n)^2} = \frac{\pi^2}{\sin^2(\pi\tau)}.$$

(c) Can one conclude that the above formula holds true whenever τ is any complex number that is not an integer?

[Hint: For (a), use residues to prove that $\hat{f}(\xi) = 0$, if $\xi < 0$, and

$$\hat{f}(\xi) = \frac{(-2\pi i)^k}{(k-1)!} \xi^{k-1} e^{2\pi i \xi \tau}, \qquad \text{when } \xi > 0.]$$

8. Suppose $\hat{f}$ has compact support contained in $[-M, M]$ and let $f(z) = \sum_{n=0}^{\infty} a_n z^n$. Show that

$$a_n = \frac{(2\pi i)^n}{n!} \int_{-M}^{M} \hat{f}(\xi) \xi^n \, d\xi,$$

and as a result

$$\limsup_{n\to\infty} (n!|a_n|)^{1/n} \leq 2\pi M.$$

In the converse direction, let f be *any* power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ with $\limsup_{n\to\infty}(n!|a_n|)^{1/n} \leq 2\pi M$. Then, f is holomorphic in the complex plane, and for every $\epsilon > 0$ there exists $A_\epsilon > 0$ such that

$$|f(z)| \leq A_\epsilon e^{2\pi(M+\epsilon)|z|}.$$

9. Here are further results similar to the Phragmén-Lindelöf theorem.

(a) Let F be a holomorphic function in the right half-plane that extends continuously to the boundary, that is, the imaginary axis. Suppose that $|F(iy)| \leq 1$ for all $y \in \mathbb{R}$, and

$$|F(z)| \leq Ce^{c|z|^\gamma}$$

for some $c, C > 0$ and $\gamma < 1$. Prove that $|F(z)| \leq 1$ for all z in the right half-plane.

(b) More generally, let S be a sector whose vertex is the origin, and forming an angle of π/β. Let F be a holomorphic function in S that is continuous on the closure of S, so that $|F(z)| \leq 1$ on the boundary of S and

$$|F(z)| \leq Ce^{c|z|^\alpha} \text{for all } z \in S$$

for some $c, C > 0$ and $0 < \alpha < \beta$. Prove that $|F(z)| \leq 1$ for all $z \in S$.

10. This exercise generalizes some of the properties of $e^{-\pi x^2}$ related to the fact that it is its own Fourier transform.

Suppose $f(z)$ is an entire function that satisfies

$$|f(x+iy)| \leq ce^{-ax^2+by^2}$$

for some $a, b, c > 0$. Let

$$\hat{f}(\zeta) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ix\zeta}\, dx.$$

Then, $\hat{f}$ is an entire function of ζ that satisfies

$$|\hat{f}(\xi + i\eta)| \leq c' e^{-a'\xi^2+b'\eta^2}$$

for some $a', b', c' > 0$.

[Hint: To prove $\hat{f}(\xi) = O(e^{-a'\xi^2})$, assume $\xi > 0$ and change the contour of integration to $x - iy$ for some $y > 0$ fixed, and $-\infty < x < \infty$. Then

$$\hat{f}(\xi) = O(e^{-2\pi y\xi}e^{by^2}).$$

Finally, choose $y = d\xi$ where d is a small constant.]

11. One can give a neater formulation of the result in Exercise 10 by proving the following fact.

Suppose $f(z)$ is an entire function of strict order 2, that is,

$$f(z) = O(e^{c_1|z|^2})$$

for some $c_1 > 0$. Suppose also that for x real,

$$f(x) = O(e^{-c_2|x|^2})$$

for some $c_2 > 0$. Then

$$|f(x+iy)| = O(e^{-ax^2+by^2})$$

for some $a, b > 0$. The converse is obviously true.

12. The principle that a function and its Fourier transform cannot both be too small at infinity is illustrated by the following theorem of Hardy.

If f is a function on $\mathbb{R}$ that satisfies

$$f(x) = O(e^{-\pi x^2}) \quad \text{and} \quad \hat{f}(\xi) = O(e^{-\pi\xi^2}),$$

then f is a constant multiple of $e^{-\pi x^2}$. As a result, if $f(x) = O(e^{-\pi Ax^2})$, and $\hat{f}(\xi) = O(e^{-\pi B\xi^2})$, with $AB > 1$ and $A, B > 0$, then f is identically zero.

(a) If f is even, show that $\hat{f}$ extends to an even entire function. Moreover, if $g(z) = \hat{f}(z^{1/2})$, then g satisfies

$$|g(x)| \leq ce^{-\pi x} \quad \text{and} \quad |g(z)| \leq ce^{\pi R\sin^2(\theta/2)} \leq ce^{\pi|z|}$$

when $x \in \mathbb{R}$ and $z = Re^{i\theta}$ with $R \geq 0$ and $\theta \in \mathbb{R}$.

(b) Apply the Phragmén-Lindelöf principle to the function

$$F(z) = g(z)e^{\gamma z} \quad \text{where } \gamma = i\pi\frac{e^{-i\pi/(2\beta)}}{\sin \pi/(2\beta)}$$

and the sector $0 \leq \theta \leq \pi/\beta < \pi$, and let $\beta \to \pi$ to deduce that $e^{\pi z}g(z)$ is bounded in the closed upper half-plane. The same result holds in the lower half-plane, so by Liouville's theorem $e^{\pi z}g(z)$ is constant, as desired.

(c) If f is odd, then $\hat{f}(0) = 0$, and apply the above argument to $\hat{f}(z)/z$ to deduce that $f = \hat{f} = 0$. Finally, write an arbitrary f as an appropriate sum of an even function and an odd function.

5 Problems

1. Suppose $\hat{f}(\xi) = O(e^{-a|\xi|^p})$ as $|\xi| \to \infty$, for some $p > 1$. Then f is holomorphic for all z and satisfies the growth condition

$$|f(z)| \leq Ae^{a|z|^q}$$

where $1/p + 1/q = 1$.

Note that on the one hand, when $p \to \infty$ then $q \to 1$, and this limiting case can be interpreted as part of Theorem 3.3. On the other hand, when $p \to 1$ then $q \to \infty$, and this limiting case in a sense brings us back to Theorem 2.1.

[Hint: To prove the result, use the inequality $-\xi^p + \xi u \leq u^q$, which is valid when ξ and u are non-negative. To establish this inequality, examine separately the cases $\xi^p \geq \xi u$ and $\xi^p < \xi u$; note also that the functions $\xi = u^{q-1}$ and $u = \xi^{p-1}$ are inverses of each other because $(p-1)(q-1) = 1$.]

2. The problem is to solve the differential equation

$$a_n \frac{d^n}{dt^n} u(t) + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} u(t) + \cdots + a_0 u(t) = f(t)\,,$$

where $a_0, a_1, \ldots, a_n$ are complex constants, and f is a given function. Here we suppose that f has bounded support and is smooth (say of class C^2).

(a) Let

$$\hat{f}(z) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i z t}\, dt.$$

Observe that $\hat{f}$ is an entire function, and using integration by parts show that

$$|\hat{f}(x+iy)| \leq \frac{A}{1+x^2}$$

if $|y| \leq a$ for any fixed $a \geq 0$.

(b) Write

$$P(z) = a_n (2\pi i z)^n + a_{n-1}(2\pi i z)^{n-1} + \cdots + a_0.$$

Find a real number c so that $P(z)$ does not vanish on the line

$$L = \{z : z = x + ic, \quad x \in \mathbb{R}\}.$$

(c) Set

$$u(t) = \int_L \frac{e^{2\pi i z t}}{P(z)} \hat{f}(z)\, dz.$$

Check that

$$\sum_{j=0}^{n} a_j \left(\frac{d}{dt}\right)^j u(t) = \int_L e^{2\pi i z t} \hat{f}(z)\, dz$$

and

$$\int_L e^{2\pi i z t} \hat{f}(z)\, dz = \int_{-\infty}^{\infty} e^{2\pi i x t} \hat{f}(x)\, dx.$$

Conclude by the Fourier inversion theorem that

$$\sum_{j=0}^{n} a_j \left(\frac{d}{dt}\right)^j u(t) = f(t).$$

Note that the solution u depends on the choice c.

3.* In this problem, we investigate the behavior of certain bounded holomorphic functions in an infinite strip. The particular result described here is sometimes called the three-lines lemma.

(a) Suppose $F(z)$ is holomorphic and bounded in the strip $0 < \text{Im}(z) < 1$ and continuous on its closure. If $|F(z)| \leq 1$ on the boundary lines, then $|F(z)| \leq 1$ throughout the strip.

(b) For the more general F, let $\sup_{x\in\mathbb{R}} |F(x)| = M_0$ and $\sup_{x\in\mathbb{R}} |F(x+i)| = M_1$. Then,

$$\sup_{x\in\mathbb{R}} |F(x+iy)| \leq M_0^{1-y} M_1^y, \qquad \text{if } 0 \leq y \leq 1.$$

(c) As a consequence, prove that $\log \sup_{x\in\mathbb{R}} |F(x+iy)|$ is a convex function of y when $0 \leq y \leq 1$.

[Hint: For part (a), apply the maximum modulus principle to $F_\epsilon(z) = F(z)e^{-\epsilon z^2}$. For part (b), consider $M_0^{z-1}M_1^{-z}F(z)$.]

4.* There is a relation between the Paley-Wiener theorem and an earlier representation due to E. Borel.

(a) A function $f(z)$, holomorphic for all z, satisfies $|f(z)| \leq A_\epsilon e^{2\pi(M+\epsilon)|z|}$ for all ϵ if and only if it is representable in the form

$$f(z) = \int_C e^{2\pi i z w} g(w)\, dw$$

where g is holomorphic outside the circle of radius M centered at the origin, and g vanishes at infinity. Here C is any circle centered at the origin of radius larger than M. In fact, if $f(z) = \sum a_n z^n$, then $g(w) = \sum_{n=0}^\infty A_n w^{-n-1}$ with $a_n = A_n (2\pi i)^{n+1}/n!$.

(b) The connection with Theorem 3.3 is as follows. For these functions f (for which in addition f and $\hat{f}$ are of moderate decrease on the real axis), one can assert that the g above is holomorphic in the larger region, which consists of the slit plane $\mathbb{C} - [-M, M]$. Moreover, the relation between g and the Fourier transform $\hat{f}$ is

$$g(z) = \frac{1}{2\pi i} \int_{-M}^{M} \frac{\hat{f}(\xi)}{\xi - z}\, d\xi$$

so that $\hat{f}$ represents the jump of g across the segment $[-M, M]$; that is,

$$\hat{f}(x) = \lim_{\epsilon\to 0, \epsilon>0} g(x+i\epsilon) - g(x-i\epsilon).$$

See Problem 5 in Chapter 3.

5 Entire Functions

> ...but after the 15th of October I felt myself a free man, with such longing for mathematical work, that the last two months flew by quickly, and that only today I found the letter of the 19th of October that I had not answered. The result of my work, with which I am not entirely satisfied, I want to share with you.
>
> Firstly, in looking back at my lectures, a gap in function theory needed to be filled. As you know, up to now the following question had been unresolved. Given an arbitrary sequence of complex numbers, $a_1, a_2, \ldots,$ can one construct an entire (transcendental) function that vanishes at these values, with prescribed multiplicities, and nowhere else?...
>
> *K. Weierstrass,* 1874

In this chapter, we will study functions that are holomorphic in the whole complex plane; these are called **entire** functions. Our presentation will be organized around the following three questions:

1. Where can such functions vanish? We shall see that the obvious necessary condition is also sufficient: if $\{z_n\}$ is any sequence of complex numbers having no limit point in $\mathbb{C}$, then there exists an entire function vanishing exactly at the points of this sequence. The construction of the desired function is inspired by Euler's product formula for $\sin \pi z$ (the prototypical case when $\{z_n\}$ is $\mathbb{Z}$), but requires an additional refinement: the Weierstrass canonical factors.

2. How do these functions grow at infinity? Here, matters are controlled by an important principle: the larger a function is, the more zeros it can have. This principle already manifests itself in the simple case of polynomials. By the fundamental theorem of algebra, the number of zeros of a polynomial P of degree d is precisely d, which is also the exponent in the order of (polynomial) growth of P, namely

$$\sup_{|z|=R} |P(z)| \approx R^d \quad \text{as } R \to \infty.$$

A precise version of this general principle is contained in Jensen's formula, which we prove in the first section. This formula, central to much of the theory developed in this chapter, exhibits a deep connection between the number of zeros of a function in a disc and the (logarithmic) average of the function over the circle. In fact, Jensen's formula not only constitutes a natural starting point for us, but also leads to the fruitful theory of value distributions, also called Nevanlinna theory (which, however, we do not take up here).

3. To what extent are these functions determined by their zeros? It turns out that if an entire function has a finite (exponential) order of growth, then it can be specified by its zeros up to multiplication by a simple factor. The precise version of this assertion is the Hadamard factorization theorem. It may be viewed as another instance of the general rule that was formulated in Chapter 3, that is, that under appropriate conditions, a holomorphic function is essentially determined by its zeros.

1 Jensen's formula

In this section, we denote by D_R and C_R the open disc and circle of radius R centered at the origin. We shall also, in the rest of this chapter, exclude the trivial case of the function that vanishes identically.

Theorem 1.1 *Let Ω be an open set that contains the closure of a disc D_R and suppose that f is holomorphic in Ω, $f(0) \neq 0$, and f vanishes nowhere on the circle C_R. If $z_1, \ldots, z_N$ denote the zeros of f inside the disc (counted with multiplicities),[1] then*

$$\log|f(0)| = \sum_{k=1}^{N} \log\left(\frac{|z_k|}{R}\right) + \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\, d\theta. \tag{1}$$

The proof of the theorem consists of several steps.

Step 1. First, we observe that if f_1 and f_2 are two functions satisfying the hypotheses and the conclusion of the theorem, then the product $f_1 f_2$ also satisfies the hypothesis of the theorem and formula (1). This observation is a simple consequence of the fact that $\log xy = \log x + \log y$ whenever x and y are positive numbers, and that the set of zeros of $f_1 f_2$ is the union of the sets of zeros of f_1 and f_2.

---

[1] That is, each zero appears in the sequence as many times as its order.

Step 2. The function

$$g(z) = \frac{f(z)}{(z - z_1) \cdots (z - z_N)}$$

initially defined on $\Omega - \{z_1, \ldots, z_N\}$, is bounded near each z_j. Therefore each z_j is a removable singularity, and hence we can write

$$f(z) = (z - z_1) \cdots (z - z_N) g(z)$$

where g is holomorphic in Ω and nowhere vanishing in the closure of D_R. By Step 1, it suffices to prove Jensen's formula for functions like g that vanish nowhere, and for functions of the form $z - z_j$.

Step 3. We first prove (1) for a function g that vanishes nowhere in the closure of D_R. More precisely, we must establish the following identity:

$$\log |g(0)| = \frac{1}{2\pi} \int_0^{2\pi} \log |g(Re^{i\theta})| \, d\theta.$$

In a slightly larger disc, we can write $g(z) = e^{h(z)}$ where h is holomorphic in that disc. This is possible since discs are simply connected, and we can define $h = \log g$ (see Theorem 6.2 in Chapter 3). Now we observe that

$$|g(z)| = |e^{h(z)}| = |e^{\mathrm{Re}(h(z)) + i \,\mathrm{Im}(h(z))}| = e^{\mathrm{Re}(h(z))},$$

so that $\log |g(z)| = \mathrm{Re}(h(z))$. The mean value property (Corollary 7.3 in Chapter 3) then immediately implies the desired formula for g.

Step 4. The last step is to prove the formula for functions of the form $f(z) = z - w$, where $w \in D_R$. That is, we must show that

$$\log |w| = \log \left(\frac{|w|}{R} \right) + \frac{1}{2\pi} \int_0^{2\pi} \log |Re^{i\theta} - w| \, d\theta.$$

Since $\log(|w|/R) = \log |w| - \log R$ and $\log |Re^{i\theta} - w| = \log R + \log |e^{i\theta} - w/R|$, it suffices to prove that

$$\int_0^{2\pi} \log |e^{i\theta} - a| \, d\theta = 0, \qquad \text{whenever } |a| < 1.$$

This in turn is equivalent (after the change of variables $\theta \mapsto -\theta$) to

$$\int_0^{2\pi} \log |1 - ae^{i\theta}| \, d\theta = 0, \qquad \text{whenever } |a| < 1.$$

To prove this, we use the function $F(z) = 1 - az$, which vanishes nowhere in the closure of the unit disc. As a consequence, there exists a holomorphic function G in a disc of radius greater than 1 such that $F(z) = e^{G(z)}$. Then $|F| = e^{\text{Re}(G)}$, and therefore $\log |F| = \text{Re}(G)$. Since $F(0) = 1$ we have $\log |F(0)| = 0$, and an application of the mean value property (Corollary 7.3 in Chapter 3) to the harmonic function $\log |F(z)|$ concludes the proof of the theorem.

From Jensen's formula we can derive an identity linking the growth of a holomorphic function with its number of zeros inside a disc. If f is a holomorphic function on the closure of a disc D_R, we denote by $\mathfrak{n}(r)$ (or $\mathfrak{n}_f(r)$ when it is necessary to keep track of the function in question) the number of zeros of f (counted with their multiplicities) inside the disc D_r, with $0 < r < R$. A simple but useful observation is that $\mathfrak{n}(r)$ is a non-decreasing function of r.

We claim that if $f(0) \neq 0$, and f does not vanish on the circle C_R, then

$$\int_0^R \mathfrak{n}(r) \, \frac{dr}{r} = \frac{1}{2\pi} \int_0^{2\pi} \log |f(Re^{i\theta})| \, d\theta - \log |f(0)|. \tag{2}$$

This formula is immediate from Jensen's equality and the next lemma.

Lemma 1.2 *If $z_1, \ldots, z_N$ are the zeros of f inside the disc D_R, then*

$$\int_0^R \mathfrak{n}(r) \, \frac{dr}{r} = \sum_{k=1}^N \log \left| \frac{R}{z_k} \right|.$$

Proof. First we have

$$\sum_{k=1}^N \log \left| \frac{R}{z_k} \right| = \sum_{k=1}^N \int_{|z_k|}^R \frac{dr}{r}.$$

If we define the characteristic function

$$\eta_k(r) = \begin{cases} 1 & \text{if } r > |z_k|, \\ 0 & \text{if } r \leq |z_k|, \end{cases}$$

then $\sum_{k=1}^N \eta_k(r) = \mathfrak{n}(r)$, and the lemma is proved using

$$\sum_{k=1}^N \int_{|z_k|}^R \frac{dr}{r} = \sum_{k=1}^N \int_0^R \eta_k(r) \, \frac{dr}{r} = \int_0^R \left(\sum_{k=1}^N \eta_k(r) \right) \frac{dr}{r} = \int_0^R \mathfrak{n}(r) \, \frac{dr}{r}.$$

2 Functions of finite order

Let f be an entire function. If there exist a positive number ρ and constants $A, B > 0$ such that

$$|f(z)| \leq Ae^{B|z|^\rho} \quad \text{for all } z \in \mathbb{C},$$

then we say that f has **an order of growth** $\leq \rho$. We define *the* **order of growth of** f as

$$\rho_f = \inf \rho\,,$$

where the infimum is over all $\rho > 0$ such that f has an order of growth $\leq \rho$.

For example, the order of growth of the function e^{z^2} is 2.

Theorem 2.1 *If f is an entire function that has an order of growth $\leq \rho$, then:*

(i) $\mathfrak{n}(r) \leq Cr^\rho$ *for some $C > 0$ and all sufficiently large r.*

(ii) *If $z_1, z_2, \ldots$ denote the zeros of f, with $z_k \neq 0$, then for all $s > \rho$ we have*

$$\sum_{k=1}^{\infty} \frac{1}{|z_k|^s} < \infty.$$

Proof. It suffices to prove the estimate for $\mathfrak{n}(r)$ when $f(0) \neq 0$. Indeed, consider the function $F(z) = f(z)/z^\ell$ where ℓ is the order of the zero of f at the origin. Then $\mathfrak{n}_f(r)$ and $\mathfrak{n}_F(r)$ differ only by a constant, and F also has an of order of growth $\leq \rho$.

If $f(0) \neq 0$ we may use formula (2), namely

$$\int_0^R \mathfrak{n}(x)\,\frac{dx}{x} = \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\,d\theta - \log|f(0)|.$$

Choosing $R = 2r$, this formula implies

$$\int_r^{2r} \mathfrak{n}(x)\,\frac{dx}{x} \leq \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\,d\theta - \log|f(0)|.$$

On the one hand, since $\mathfrak{n}(r)$ is increasing, we have

$$\int_r^{2r} \mathfrak{n}(x)\,\frac{dx}{x} \geq \mathfrak{n}(r)\int_r^{2r} \frac{dx}{x} = \mathfrak{n}(r)[\log 2r - \log r] = \mathfrak{n}(r)\log 2,$$

and on the other hand, the growth condition on f gives

$$\int_0^{2\pi} \log |f(Re^{i\theta})|\, d\theta \leq \int_0^{2\pi} \log |Ae^{BR^\rho}|\, d\theta \leq C' r^\rho$$

for all large r. Consequently, $\mathfrak{n}(r) \leq Cr^\rho$ for an appropriate $C > 0$ and all sufficiently large r.

The following estimates prove the second part of the theorem:

$$\begin{aligned}
\sum_{|z_k|\geq 1} |z_k|^{-s} &= \sum_{j=0}^\infty \left(\sum_{2^j \leq |z_k| < 2^{j+1}} |z_k|^{-s} \right) \\
&\leq \sum_{j=0}^\infty 2^{-js} \mathfrak{n}(2^{j+1}) \\
&\leq c \sum_{j=0}^\infty 2^{-js} 2^{(j+1)\rho} \\
&\leq c' \sum_{j=0}^\infty (2^{\rho - s})^j \\
&< \infty.
\end{aligned}$$

The last series converges because $s > \rho$.

Part (ii) of the theorem is a noteworthy fact, which we shall use in a later part of this chapter.

We give two simple examples of the theorem; each of these shows that the condition $s > \rho$ cannot be improved.

Example 1. Consider $f(z) = \sin \pi z$. Recall Euler's identity, namely

$$f(z) = \frac{e^{i\pi z} - e^{-i\pi z}}{2i},$$

which implies that $|f(z)| \leq e^{\pi |z|}$, and f has an order of growth ≤ 1. By taking $z = ix$, where $x \in \mathbb{R}$, it is clear that the order of growth of f is actually equal to 1. However, f vanishes to order 1 at $z = n$ for each $n \in \mathbb{Z}$, and $\sum_{n\neq 0} 1/|n|^s < \infty$ precisely when $s > 1$.

Example 2. Consider $f(z) = \cos z^{1/2}$, which we *define* by

$$\cos z^{1/2} = \sum_{n=0}^\infty (-1)^n \frac{z^n}{(2n)!}.$$

Then f is entire, and it is easy to see that

$$|f(z)| \leq e^{|z|^{1/2}},$$

and the order of growth of f is 1/2. Moreover, $f(z)$ vanishes when $z_n = ((n + 1/2)\pi)^2$, while $\sum_n 1/|z_n|^s < \infty$ exactly when $s > 1/2$.

A natural question is whether or not, given any sequence of complex numbers $z_1, z_2, \ldots$, there exists an entire function f with zeros precisely at the points of this sequence. A necessary condition is that $z_1, z_2, \ldots$ do not accumulate, in other words we must have

$$\lim_{k \to \infty} |z_k| = \infty\,,$$

otherwise f would vanish identically by Theorem 4.8 in Chapter 2. Weierstrass proved that this condition is also sufficient by explicitly constructing a function with these prescribed zeros. A first guess is of course the product

$$(z - z_1)(z - z_2) \cdots\,,$$

which provides a solution in the special case when the sequence of zeros is finite. In general, Weierstrass showed how to insert factors in this product so that the convergence is guaranteed, yet no new zeros are introduced.

Before coming to the general construction, we review infinite products and study a basic example.

3 Infinite products

3.1 Generalities

Given a sequence $\{a_n\}_{n=1}^{\infty}$ of complex numbers, we say that the product

$$\prod_{n=1}^{\infty}(1 + a_n)$$

converges if the limit

$$\lim_{N \to \infty} \prod_{n=1}^{N}(1 + a_n)$$

of the partial products exists.

A useful necessary condition that guarantees the existence of a product is contained in the following proposition.

Proposition 3.1 *If* $\sum |a_n| < \infty$*, then the product* $\prod_{n=1}^{\infty}(1+a_n)$ *converges. Moreover, the product converges to* 0 *if and only if one of its factors is* 0*.*

This is simply Proposition 1.9 of Chapter 8 in Book I. We repeat the proof here.

Proof. If $\sum |a_n|$ converges, then for all large n we must have $|a_n| < 1/2$. Disregarding if necessary finitely many terms, we may assume that this inequality holds for all n. In particular, we can define $\log(1+a_n)$ by the usual power series (see (6) in Chapter 3), and this logarithm satisfies the property that $1+z = e^{\log(1+z)}$ whenever $|z|<1$. Hence we may write the partial products as follows:

$$\prod_{n=1}^{N}(1+a_n) = \prod_{n=1}^{N} e^{\log(1+a_n)} = e^{B_N}\,,$$

where $B_N = \sum_{n=1}^{N} b_n$ with $b_n = \log(1+a_n)$. By the power series expansion we see that $|\log(1+z)| \leq 2|z|$, if $|z| < 1/2$. Hence $|b_n| \leq 2|a_n|$, so B_N converges as $N \to \infty$ to a complex number, say B. Since the exponential function is continuous, we conclude that e^{B_N} converges to e^B as $N \to \infty$, proving the first assertion of the proposition. Observe also that if $1 + a_n \neq 0$ for all n, then the product converges to a non-zero limit since it is expressed as e^B.

More generally, we can consider products of holomorphic functions.

Proposition 3.2 *Suppose* $\{F_n\}$ *is a sequence of holomorphic functions on the open set* Ω*. If there exist constants* $c_n > 0$ *such that*

$$\sum c_n < \infty \quad \text{and} \quad |F_n(z) - 1| \leq c_n \quad \text{for all } z \in \Omega,$$

then:

(i) *The product* $\prod_{n=1}^{\infty} F_n(z)$ *converges uniformly in* Ω *to a holomorphic function* $F(z)$*.*

(ii) *If* $F_n(z)$ *does not vanish for any* n*, then*

$$\frac{F'(z)}{F(z)} = \sum_{n=1}^{\infty} \frac{F_n'(z)}{F_n(z)}.$$

Proof. To prove the first statement, note that for each z we may argue as in the previous proposition if we write $F_n(z) = 1 + a_n(z)$, with

$|a_n(z)| \leq c_n$. Then, we observe that the estimates are actually uniform in z because the c_n's are constants. It follows that the product converges uniformly to a holomorphic function, which we denote by $F(z)$.

To establish the second part of the theorem, suppose that K is a compact subset of Ω, and let

$$G_N(z) = \prod_{n=1}^{N} F_n(z).$$

We have just proved that $G_N \to F$ uniformly in Ω, so by Theorem 5.3 in Chapter 2, the sequence $\{G'_N\}$ converges uniformly to F' in K. Since G_N is uniformly bounded from below on K, we conclude that $G'_N/G_N \to F'/F$ uniformly on K, and because K is an arbitrary compact subset of Ω, the limit holds for every point of Ω. Moreover, as we saw in Section 4 of Chapter 3

$$\frac{G'_N}{G_N} = \sum_{n=1}^{N} \frac{F'_n}{F_n},$$

so part (ii) of the proposition is also proved.

3.2 Example: the product formula for the sine function

Before proceeding with the general theory of Weierstrass products, we consider the key example of the product formula for the sine function:

$$\frac{\sin \pi z}{\pi} = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right). \tag{3}$$

This identity will in turn be derived from the sum formula for the cotangent function ($\cot \pi z = \cos \pi z / \sin \pi z$):

$$\pi \cot \pi z = \sum_{n=-\infty}^{\infty} \frac{1}{z+n} = \lim_{N\to\infty} \sum_{|n|\leq N} \frac{1}{z+n} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}. \tag{4}$$

The first formula holds for all complex numbers z, and the second whenever z is not an integer. The sum $\sum_{n=-\infty}^{\infty} 1/(z+n)$ needs to be properly understood, because the separate halves corresponding to positive and negative values of n do not converge. Only when interpreted symmetrically, as $\lim_{N\to\infty} \sum_{|n|\leq N} 1/(z+n)$, does the cancellation of terms lead to a convergent series as in (4) above.

We prove (4) by showing that both $\pi \cot \pi z$ and the series have the same structural properties. In fact, observe that if $F(z) = \pi \cot \pi z$, then F has the following three properties:

(i) $F(z+1) = F(z)$ whenever z is not an integer.

(ii) $F(z) = \dfrac{1}{z} + F_0(z)$, where F_0 is analytic near 0.

(iii) $F(z)$ has simple poles at the integers, and no other singularities.

Then, we note that the function

$$\sum_{n=-\infty}^{\infty} \frac{1}{z+n} = \lim_{N\to\infty} \sum_{|n|\le N} \frac{1}{z+n}$$

also satisfies these same three properties. In fact, property (i) is simply the observation that the passage from z to $z+1$ merely shifts the terms in the infinite sum. To be precise,

$$\sum_{|n|\le N} \frac{1}{z+1+n} = \frac{1}{z+1+N} - \frac{1}{z-N} + \sum_{|n|\le N} \frac{1}{z+n}.$$

Letting N tend to infinity proves the assertion. Properties (ii) and (iii) are evident from the representation $\frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2-n^2}$ of the sum.

Therefore, the function defined by

$$\Delta(z) = F(z) - \sum_{n=-\infty}^{\infty} \frac{1}{z+n}$$

is periodic in the sense that $\Delta(z+1) = \Delta(z)$, and by (ii) the singularity of Δ at the origin is removable, and hence by periodicity the singularities at all the integers are also removable; this implies that Δ is entire.

To prove our formula, it will suffice to show that the function Δ is bounded in the complex plane. By the periodicity above, it is enough to do so in the strip $|\mathrm{Re}(z)| \le 1/2$. This is because every $z' \in \mathbb{C}$ is of the form $z' = z + k$, where z is in the strip and k is an integer. Since Δ is holomorphic, it is bounded in the rectangle $|\mathrm{Im}(z)| \le 1$, and we need only control the behavior of that function for $|\mathrm{Im}(z)| > 1$. If $\mathrm{Im}(z) > 1$ and $z = x + iy$, then

$$\cot \pi z = i\,\frac{e^{i\pi z} + e^{-i\pi z}}{e^{i\pi z} - e^{-i\pi z}} = i\,\frac{e^{-2\pi y} + e^{-2\pi i x}}{e^{-2\pi y} - e^{-2\pi i x}},$$

and in absolute value this quantity is bounded. Also

$$\frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2} = \frac{1}{x+iy} + \sum_{n=1}^{\infty} \frac{2(x+iy)}{x^2 - y^2 - n^2 + 2ixy};$$

therefore if $y > 1$, we have

$$\left| \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2} \right| \leq C + C \sum_{n=1}^{\infty} \frac{y}{y^2 + n^2}.$$

Now the sum on the right-hand side is majorized by

$$\int_0^{\infty} \frac{y}{y^2 + x^2}\, dx\,,$$

because the function $y/(y^2 + x^2)$ is decreasing in x; moreover, as the change of variables $x \mapsto yx$ shows, the integral is independent of y, and hence bounded. By a similar argument Δ is bounded in the strip where $\mathrm{Im}(z) < -1$, hence is bounded throughout the whole strip $|\mathrm{Re}(z)| \leq 1/2$. Therefore Δ is bounded in $\mathbb{C}$, and by Liouville's theorem, $\Delta(z)$ is constant. The observation that Δ is odd shows that this constant must be 0, and concludes the proof of formula (4).

To prove (3), we now let

$$G(z) = \frac{\sin \pi z}{\pi} \qquad \text{and} \qquad P(z) = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$

Proposition 3.2 and the fact that $\sum 1/n^2 < \infty$ guarantee that the product $P(z)$ converges, and that away from the integers we have

$$\frac{P'(z)}{P(z)} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}.$$

Since $G'(z)/G(z) = \pi \cot \pi z$, the cotangent formula (4) gives

$$\left(\frac{P(z)}{G(z)}\right)' = \frac{P(z)}{G(z)} \left[\frac{P'(z)}{P(z)} - \frac{G'(z)}{G(z)}\right] = 0,$$

and so $P(z) = cG(z)$ for some constant c. Dividing this identity by z, and taking the limit as $z \to 0$, we find $c = 1$.

Remark. Other proofs of (4) and (3) can be given by integrating analogous identities for $\pi^2/(\sin \pi z)^2$ derived in Exercise 12, Chapter 3, and Exercise 7, Chapter 4. Still other proofs using Fourier series can be found in the exercises of Chapters 3 and 5 of Book I.

4 Weierstrass infinite products

We now turn to Weierstrass's construction of an entire function with prescribed zeros.

Theorem 4.1 *Given any sequence $\{a_n\}$ of complex numbers with $|a_n| \to \infty$ as $n \to \infty$, there exists an entire function f that vanishes at all $z = a_n$ and nowhere else. Any other such entire function is of the form $f(z)e^{g(z)}$, where g is entire.*

Recall that if a holomorphic function f vanishes at $z = a$, then the multiplicity of the zero a is the integer m so that

$$f(z) = (z - a)^m g(z),$$

where g is holomorphic and nowhere vanishing in a neighborhood of a. Alternatively, m is the first non-zero power of $z - a$ in the power series expansion of f at a. Since, as before, we allow for repetitions in the sequence $\{a_n\}$, the theorem actually guarantees the existence of entire functions with prescribed zeros and with desired multiplicities.

To begin the proof, note first that if f_1 and f_2 are two entire functions that vanish at all $z = a_n$ and nowhere else, then f_1/f_2 has removable singularities at all the points a_n. Hence f_1/f_2 is entire and vanishes nowhere, so that there exists an entire function g with $f_1(z)/f_2(z) = e^{g(z)}$, as we showed in Section 6 of Chapter 3. Therefore $f_1(z) = f_2(z)e^{g(z)}$ and the last statement of the theorem is verified.

Hence we are left with the task of constructing a function that vanishes at all the points of the sequence $\{a_n\}$ and nowhere else. A naive guess, suggested by the product formula for $\sin \pi z$, is the product $\prod_n (1 - z/a_n)$. The problem is that this product converges only for suitable sequences $\{a_n\}$, so we correct this by inserting exponential factors. These factors will make the product converge without adding new zeros.

For each integer $k \geq 0$ we define **canonical factors** by

$$E_0(z) = 1 - z \quad \text{and} \quad E_k(z) = (1 - z)e^{z + z^2/2 + \cdots + z^k/k}, \quad \text{for } k \geq 1.$$

The integer k is called the **degree** of the canonical factor.

Lemma 4.2 *If $|z| \leq 1/2$, then $|1 - E_k(z)| \leq c|z|^{k+1}$ for some $c > 0$.*

Proof. If $|z| \leq 1/2$, then with the logarithm defined in terms of the power series, we have $1 - z = e^{\log(1-z)}$, and therefore

$$E_k(z) = e^{\log(1-z) + z + z^2/2 + \cdots + z^k/k} = e^w,$$

where $w = -\sum_{n=k+1}^{\infty} z^n/n$. Observe that since $|z| \leq 1/2$ we have

$$|w| \leq |z|^{k+1} \sum_{n=k+1}^{\infty} |z|^{n-k-1}/n \leq |z|^{k+1} \sum_{j=0}^{\infty} 2^{-j} \leq 2|z|^{k+1}.$$

In particular, we have $|w| \leq 1$ and this implies that

$$|1 - E_k(z)| = |1 - e^w| \leq c'|w| \leq c|z|^{k+1}.$$

Remark. An important technical point is that the constant c in the statement of the lemma can be chosen to be independent of k. In fact, an examination of the proof shows that we may take $c' = e$ and then $c = 2e$.

Suppose that we are given a zero of order m at the origin, and that $a_1, a_2 \dots$ are all non-zero. Then we define the Weierstrass product by

$$f(z) = z^m \prod_{n=1}^{\infty} E_n(z/a_n).$$

We claim that this function has the required properties; that is, f is entire with a zero of order m at the origin, zeros at each point of the sequence $\{a_n\}$, and f vanishes nowhere else.

Fix $R > 0$, and suppose that z belongs to the disc $|z| < R$. We shall prove that f has all the desired properties in this disc, and since R is arbitrary, this will prove the theorem.

We can consider two types of factors in the formula defining f, with the choice depending on whether $|a_n| \leq 2R$ or $|a_n| > 2R$. There are only finitely many terms of the first kind (since $|a_n| \to \infty$), and we see that the finite product vanishes at all $z = a_n$ with $|a_n| < R$. If $|a_n| \geq 2R$, we have $|z/a_n| \leq 1/2$, hence the previous lemma implies

$$|1 - E_n(z/a_n)| \leq c \left| \frac{z}{a_n} \right|^{n+1} \leq \frac{c}{2^{n+1}}.$$

Note that by the above remark, c does not depend on n. Therefore, the product

$$\prod_{|a_n| \geq 2R} E_n(z/a_n)$$

defines a holomorphic function when $|z| < R$, and does not vanish in that disc by the propositions in Section 3. This shows that the function

f has the desired properties, and the proof of Weierstrass's theorem is complete.

5 Hadamard's factorization theorem

The theorem of this section combines the results relating the growth of a function to the number of zeros it possesses, and the above product theorem. Weierstrass's theorem states that a function that vanishes at the points $a_1, a_2, \ldots$ takes the form

$$e^{g(z)} z^m \prod_{n=1}^{\infty} E_n(z/a_n).$$

Hadamard refined this result by showing that in the case of functions of finite order, the degree of the canonical factors can be taken to be constant, and g is then a polynomial.

Recall that an entire function has an order of growth $\leq \rho$ if

$$|f(z)| \leq Ae^{B|z|^\rho},$$

and that *the* order of growth ρ_0 of f is the infimum of all such ρ's.

A basic result we proved earlier was that if f has order of growth $\leq \rho$, then

$$\mathfrak{n}(r) \leq Cr^\rho, \qquad \text{for all large } r,$$

and if $a_1, a_2, \ldots$ are the non-zero zeros of f, and $s > \rho$, then

$$\sum |a_n|^{-s} < \infty.$$

Theorem 5.1 *Suppose f is entire and has growth order ρ_0. Let k be the integer so that $k \leq \rho_0 < k+1$. If $a_1, a_2, \ldots$ denote the (non-zero) zeros of f, then*

$$f(z) = e^{P(z)} z^m \prod_{n=1}^{\infty} E_k(z/a_n),$$

where P is a polynomial of degree $\leq k$, and m is the order of the zero of f at $z = 0$.

Main lemmas

Here we gather a few lemmas needed in the proof of Hadamard's theorem.

Lemma 5.2 *The canonical products satisfy*

$$|E_k(z)| \geq e^{-c|z|^{k+1}} \qquad \text{if } |z| \leq 1/2$$

and

$$|E_k(z)| \geq |1 - z|\, e^{-c'|z|^k} \qquad \text{if } |z| \geq 1/2.$$

Proof. If $|z| \leq 1/2$ we can use the power series to define the logarithm of $1 - z$, so that

$$E_k(z) = e^{\log(1-z)+\sum_{n=1}^{k} z^n/n} = e^{-\sum_{n=k+1}^{\infty} z^n/n} = e^w.$$

Since $|e^w| \geq e^{-|w|}$ and $|w| \leq c|z|^{k+1}$, the first part of the lemma follows. For the second part, simply observe that if $|z| \geq 1/2$, then

$$|E_k(z)| = |1 - z||e^{z+z^2/2+\cdots+z^k/k}|,$$

and that there exists $c' > 0$ such that

$$|e^{z+z^2/2+\cdots+z^k/k}| \geq e^{-|z+z^2/2+\cdots+z^k/k|} \geq e^{-c'|z|^k}.$$

The inequality in the lemma when $|z| \geq 1/2$ then follows from these observations.

The key to the proof of Hadamard's theorem consists of finding a lower bound for the product of the canonical factors when z stays away from the zeros $\{a_n\}$. Therefore, we shall first estimate the product from below, in the complement of small discs centered at these points.

Lemma 5.3 *For any s with $\rho_0 < s < k + 1$, we have*

$$\left|\prod_{n=1}^{\infty} E_k(z/a_n)\right| \geq e^{-c|z|^s},$$

except possibly when z belongs to the union of the discs centered at a_n of radius $|a_n|^{-k-1}$, for $n = 1, 2, 3, \ldots$.

Proof. The proof this lemma is a little subtle. First, we write

$$\prod_{n=1}^{\infty} E_k(z/a_n) = \prod_{|a_n| \leq 2|z|} E_k(z/a_n) \prod_{|a_n| > 2|z|} E_k(z/a_n).$$

For the second product the estimate asserted above holds with no restriction on z. Indeed, by the previous lemma

$$\begin{aligned}\left|\prod_{|a_n|>2|z|} E_k(z/a_n)\right| &= \prod_{|a_n|>2|z|} |E_k(z/a_n)| \\ &\geq \prod_{|a_n|>2|z|} e^{-c|z/a_n|^{k+1}} \\ &\geq e^{-c|z|^{k+1}\sum_{|a_n|>2|z|}|a_n|^{-k-1}}.\end{aligned}$$

But $|a_n| > 2|z|$ and $s < k+1$, so we must have

$$|a_n|^{-k-1} = |a_n|^{-s}|a_n|^{s-k-1} \leq C|a_n|^{-s}|z|^{s-k-1}.$$

Therefore, the fact that $\sum |a_n|^{-s}$ converges implies that

$$\left|\prod_{|a_n|>2|z|} E_k(z/a_n)\right| \geq e^{-c|z|^s}$$

for some $c > 0$.

To estimate the first product, we use the second part of Lemma 5.2, and write

$$\left|\prod_{|a_n|\leq 2|z|} E_k(z/a_n)\right| \geq \prod_{|a_n|\leq 2|z|} \left|1 - \frac{z}{a_n}\right| \prod_{|a_n|\leq 2|z|} e^{-c'|z/a_n|^k}. \tag{5}$$

We now note that

$$\prod_{|a_n|\leq 2|z|} e^{-c'|z/a_n|^k} = e^{-c'|z|^k\sum_{|a_n|\leq 2|z|}|a_n|^{-k}},$$

and again, we have $|a_n|^{-k} = |a_n|^{-s}|a_n|^{s-k} \leq C|a_n|^{-s}|z|^{s-k}$, thereby proving that

$$\prod_{|a_n|\leq 2|z|} e^{-c'|z/a_n|^k} \geq e^{-c|z|^s}.$$

It is the estimate on the first product on the right-hand side of (5) which requires the restriction on z imposed in the statement of the

lemma. Indeed, whenever z does not belong to a disc of radius $|a_n|^{-k-1}$ centered at a_n, we must have $|a_n - z| \geq |a_n|^{-k-1}$. Therefore

$$\begin{aligned}\prod_{|a_n|\leq 2|z|} \left|1 - \frac{z}{a_n}\right| &= \prod_{|a_n|\leq 2|z|} \left|\frac{a_n - z}{a_n}\right| \\ &\geq \prod_{|a_n|\leq 2|z|} |a_n|^{-k-1}|a_n|^{-1} \\ &= \prod_{|a_n|\leq 2|z|} |a_n|^{-k-2}.\end{aligned}$$

Finally, the estimate for the first product follows from the fact that

$$\begin{aligned}(k+2) \sum_{|a_n|\leq 2|z|} \log|a_n| &\leq (k+2)\mathfrak{n}(2|z|)\log 2|z| \\ &\leq c|z|^s \log 2|z| \\ &\leq c'|z|^{s'},\end{aligned}$$

for any $s' > s$, and the second inequality follows because $\mathfrak{n}(2|z|) \leq c|z|^s$ by Theorem 2.1. Since we restricted s to satisfy $s > \rho_0$, we can take an initial s sufficiently close to ρ_0, so that the assertion of the lemma is established (with s being replaced by s').

Corollary 5.4 *There exists a sequence of radii, $r_1, r_2, \ldots$, with $r_m \to \infty$, such that*

$$\left|\prod_{n=1}^{\infty} E_k(z/a_n)\right| \geq e^{-c|z|^s} \qquad \textit{for } |z| = r_m.$$

Proof. Since $\sum |a_n|^{-k-1} < \infty$, there exists an integer N so that

$$\sum_{n=N}^{\infty} |a_n|^{-k-1} < 1/10.$$

Therefore, given any two consecutive large integers L and $L+1$, we can find a positive number r with $L \leq r \leq L+1$, such that the circle of radius r centered at the origin does not intersect the forbidden discs of Lemma 5.3. For otherwise, the union of the intervals

$$I_n = \left[|a_n| - \frac{1}{|a_n|^{k+1}},\ |a_n| + \frac{1}{|a_n|^{k+1}}\right]$$

(which are of length $2|a_n|^{-k-1}$) would cover all the interval $[L, L+1]$. (See Figure 1.) This would imply $2\sum_{n=N}^{\infty} |a_n|^{-k-1} \geq 1$, which is a contradiction. We can then apply the previous lemma with $|z| = r$ to conclude the proof of the corollary.

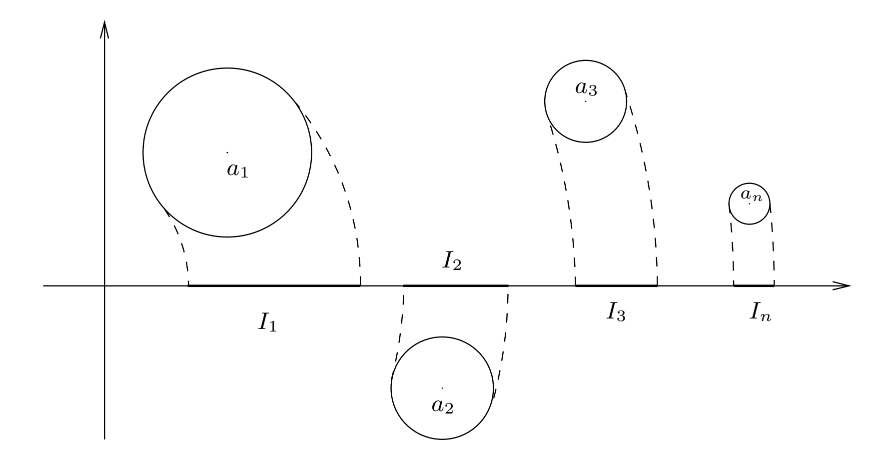


Figure 1. The intervals I_n

Proof of Hadamard's theorem

Let

$$E(z) = z^m \prod_{n=1}^{\infty} E_k(z/a_n).$$

To prove that E is entire, we repeat the argument in the proof of Theorem 4.1; we take into account that by Lemma 4.2

$$|1 - E_k(z/a_n)| \leq c \left| \frac{z}{a_n} \right|^{k+1} \qquad \text{for all large } n,$$

and that the series $\sum |a_n|^{-k-1}$ converges. (Recall $\rho_0 < s < k+1$.) Moreover, E has the zeros of f, therefore f/E is holomorphic and nowhere vanishing. Hence

$$\frac{f(z)}{E(z)} = e^{g(z)}$$

for some entire function g. By the fact that f has growth order ρ_0, and because of the estimate from below for E obtained in Corollary 5.4, we have

$$e^{\mathrm{Re}(g(z))} = \left|\frac{f(z)}{E(z)}\right| \leq c' e^{c|z|^s}$$

for $|z| = r_m$. This proves that

$$\mathrm{Re}(g(z)) \leq C|z|^s, \qquad \text{for } |z| = r_m.$$

The proof of Hadamard's theorem is therefore complete if we can establish the following final lemma.

Lemma 5.5 *Suppose g is entire and $u = \mathrm{Re}(g)$ satisfies*

$$u(z) \leq Cr^s \qquad \textit{whenever } |z| = r$$

for a sequence of positive real numbers r that tends to infinity. Then g is a polynomial of degree $\leq s$.

Proof. We can expand g in a power series centered at the origin

$$g(z) = \sum_{n=0}^{\infty} a_n z^n.$$

We have already proved in the last section of Chapter 3 (as a simple application of Cauchy's integral formulas) that

$$\frac{1}{2\pi}\int_0^{2\pi} g(re^{i\theta})e^{-in\theta}\,d\theta = \begin{cases} a_n r^n & \text{if } n \geq 0 \\ 0 & \text{if } n < 0. \end{cases} \tag{6}$$

By taking complex conjugates we find that

$$\frac{1}{2\pi}\int_0^{2\pi} \overline{g(re^{i\theta})}e^{-in\theta}\,d\theta = 0 \tag{7}$$

whenever $n > 0$, and since $2u = g + \overline{g}$ we add equations (6) and (7) to obtain

$$a_n r^n = \frac{1}{\pi}\int_0^{2\pi} u(re^{i\theta})e^{-in\theta}\,d\theta, \qquad \text{whenever } n > 0.$$

For $n = 0$ we can simply take real parts of both sides of (6) to find that

$$2\mathrm{Re}(a_0) = \frac{1}{\pi}\int_0^{2\pi} u(re^{i\theta})\,d\theta.$$

Now we recall the simple fact that whenever $n \neq 0$, the integral of $e^{-in\theta}$ over any circle centered at the origin vanishes. Therefore

$$a_n = \frac{1}{\pi r^n} \int_0^{2\pi} [u(re^{i\theta}) - Cr^s] e^{-in\theta}\, d\theta \qquad \text{when } n > 0,$$

hence

$$|a_n| \leq \frac{1}{\pi r^n} \int_0^{2\pi} [Cr^s - u(re^{i\theta})]\, d\theta \leq 2Cr^{s-n} - 2\mathrm{Re}(a_0) r^{-n}.$$

Letting r tend to infinity along the sequence given in the hypothesis of the lemma proves that $a_n = 0$ for $n > s$. This completes the proof of the lemma and of Hadamard's theorem.

6 Exercises

1. Give another proof of Jensen's formula in the unit disc using the functions (called Blaschke factors)

$$\psi_\alpha(z) = \frac{\alpha - z}{1 - \overline{\alpha} z}.$$

[Hint: The function $f/(\psi_{z_1} \cdots \psi_{z_N})$ is nowhere vanishing.]

2. Find the order of growth of the following entire functions:

(a) $p(z)$ where p is a polynomial.

(b) e^{bz^n} for $b \neq 0$.

(c) e^{e^z}.

3. Show that if τ is fixed with $\mathrm{Im}(\tau) > 0$, then the Jacobi theta function

$$\Theta(z|\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{2\pi i n z}$$

is of order 2 as a function of z. Further properties of Θ will be studied in Chapter 10.
[Hint: $-n^2 t + 2n|z| \leq -n^2 t/2$ when $t > 0$ and $n \geq 4|z|/t$.]

4. Let $t > 0$ be given and fixed, and define $F(z)$ by

$$F(z) = \prod_{n=1}^{\infty} (1 - e^{-2\pi n t} e^{2\pi i z}).$$

Note that the product defines an entire function of z.

(a) Show that $|F(z)| \leq Ae^{a|z|^2}$, hence F is of order 2.

(b) F vanishes exactly when $z = -int + m$ for $n \geq 1$ and n, m integers. Thus, if z_n is an enumeration of these zeros we have

$$\sum \frac{1}{|z_n|^2} = \infty \quad \text{but} \quad \sum \frac{1}{|z_n|^{2+\epsilon}} < \infty.$$

[Hint: To prove (a), write $F(z) = F_1(z)F_2(z)$ where

$$F_1(z) = \prod_{n=1}^{N}(1 - e^{-2\pi nt}e^{2\pi iz}) \quad \text{and} \quad F_2(z) = \prod_{n=N+1}^{\infty}(1 - e^{-2\pi nt}e^{2\pi iz}).$$

Choose $N \approx c|z|$ with c appropriately large. Then, since

$$\left(\sum_{N+1}^{\infty} e^{-2\pi nt}\right) e^{2\pi|z|} \leq 1\,,$$

one has $|F_2(z)| \leq A$. However,

$$|1 - e^{-2\pi nt}e^{2\pi iz}| \leq 1 + e^{2\pi|z|} \leq 2e^{2\pi|z|}.$$

Thus $|F_1(z)| \leq 2^N e^{2\pi N|z|} \leq e^{c'|z|^2}$. Note that a simple variant of the function F arises as a factor in the triple product formula for the Jacobi theta function Θ, taken up in Chapter 10.]

5. Show that if $\alpha > 1$, then

$$F_\alpha(z) = \int_{-\infty}^{\infty} e^{-|t|^\alpha} e^{2\pi izt}\, dt$$

is an entire function of growth order $\alpha/(\alpha - 1)$.

[Hint: Show that

$$-\frac{|t|^\alpha}{2} + 2\pi|z||t| \leq c|z|^{\alpha/(\alpha-1)}$$

by considering the two cases $|t|^{\alpha-1} \leq A|z|$ and $|t|^{\alpha-1} \geq A|z|$, for an appropriate constant A.]

6. Prove Wallis's product formula

$$\frac{\pi}{2} = \frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{2m\cdot 2m}{(2m-1)\cdot(2m+1)}\cdots.$$

[Hint: Use the product formula for $\sin z$ at $z = \pi/2$.]

7. Establish the following properties of infinite products.

(a) Show that if $\sum |a_n|^2$ converges, then the product $\prod(1+a_n)$ converges to a non-zero limit if and only if $\sum a_n$ converges.

(b) Find an example of a sequence of complex numbers $\{a_n\}$ such that $\sum a_n$ converges but $\prod(1+a_n)$ diverges.

(c) Also find an example such that $\prod(1+a_n)$ converges and $\sum a_n$ diverges.

8. Prove that for every z the product below converges, and

$$\cos(z/2)\cos(z/4)\cos(z/8)\cdots = \prod_{k=1}^{\infty}\cos(z/2^k) = \frac{\sin z}{z}.$$

[Hint: Use the fact that $\sin 2z = 2\sin z\cos z$.]

9. Prove that if $|z|<1$, then

$$(1+z)(1+z^2)(1+z^4)(1+z^8)\cdots = \prod_{k=0}^{\infty}(1+z^{2^k}) = \frac{1}{1-z}.$$

10. Find the Hadamard products for:

(a) $e^z - 1$;

(b) $\cos \pi z$.

[Hint: The answers are $e^{z/2} z \prod_{n=1}^{\infty}(1+z^2/4n^2\pi^2)$ and $\prod_{n=0}^{\infty}(1-4z^2/(2n+1)^2)$, respectively.]

11. Show that if f is an entire function of finite order that omits two values, then f is constant. This result remains true for any entire function and is known as Picard's little theorem.

[Hint: If f misses a, then $f(z)-a$ is of the form $e^{p(z)}$ where p is a polynomial.]

12. Suppose f is entire and never vanishes, and that none of the higher derivatives of f ever vanish. Prove that if f is also of finite order, then $f(z) = e^{az+b}$ for some constants a and b.

13. Show that the equation $e^z - z = 0$ has infinitely many solutions in $\mathbb{C}$.

[Hint: Apply Hadamard's theorem.]

14. Deduce from Hadamard's theorem that if F is entire and of growth order ρ that is non-integral, then F has infinitely many zeros.

15. Prove that every meromorphic function in $\mathbb{C}$ is the quotient of two entire functions. Also, if $\{a_n\}$ and $\{b_n\}$ are two disjoint sequences having no finite limit

points, then there exists a meromorphic function in the whole complex plane that vanishes exactly at $\{a_n\}$ and has poles exactly at $\{b_n\}$.

16. Suppose that

$$Q_n(z) = \sum_{k=1}^{N_n} c_k^n \, z^k$$

are given polynomials for $n = 1, 2, \ldots$. Suppose also that we are given a sequence of complex numbers $\{a_n\}$ without limit points. Prove that there exists a meromorphic function $f(z)$ whose only poles are at $\{a_n\}$, and so that for each n, the difference

$$f(z) - Q_n\left(\frac{1}{z - a_n}\right)$$

is holomorphic near a_n. In other words, f has a prescribed poles and principal parts at each of these poles. This result is due to Mittag-Leffler.

17. Given two countably infinite sequences of complex numbers $\{a_k\}_{k=0}^{\infty}$ and $\{b_k\}_{k=0}^{\infty}$, with $\lim_{k\to\infty} |a_k| = \infty$, it is always possible to find an entire function F that satisfies $F(a_k) = b_k$ for all k.

(a) Given n distinct complex numbers $a_1, \ldots, a_n$, and another n complex numbers $b_1, \ldots, b_n$, construct a polynomial P of degree $\leq n - 1$ with

$$P(a_i) = b_i \qquad \text{for } i = 1, \ldots, n.$$

(b) Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of distinct complex numbers such that $a_0 = 0$ and $\lim_{k\to\infty} |a_k| = \infty$, and let $E(z)$ denote a Weierstrass product associated with $\{a_k\}$. Given complex numbers $\{b_k\}_{k=0}^{\infty}$, show that there exist integers $m_k \geq 1$ such that the series

$$F(z) = \frac{b_0}{E'(z)} \frac{E(z)}{z} + \sum_{k=1}^{\infty} \frac{b_k}{E'(a_k)} \frac{E(z)}{z - a_k} \left(\frac{z}{a_k}\right)^{m_k}$$

defines an entire function that satisfies

$$F(a_k) = b_k \qquad \text{for all } k \geq 0.$$

This is known as the Pringsheim interpolation formula.

7 Problems

1. Prove that if f is holomorphic in the unit disc, bounded and not identically zero, and $z_1, z_2, \ldots, z_n, \ldots$ are its zeros ($|z_k| < 1$), then

$$\sum_n (1 - |z_n|) < \infty.$$

[Hint: Use Jensen's formula.]

2.* In this problem, we discuss Blaschke products, which are bounded analogues in the disc of the Weierstrass products for entire functions.

(a) Show that for $0 < |\alpha| < 1$ and $|z| \leq r < 1$ the inequality

$$\left| \frac{\alpha + |\alpha| z}{(1 - \overline{\alpha} z)\alpha} \right| \leq \frac{1+r}{1-r}$$

holds.

(b) Let $\{\alpha_n\}$ be a sequence in the unit disc such that $\alpha_n \neq 0$ for all n and

$$\sum_{n=1}^{\infty} (1 - |\alpha_n|) < \infty.$$

Note that this will be the case if $\{\alpha_n\}$ are the zeros of a bounded holomorphic function on the unit disc (see Problem 1). Show that the product

$$f(z) = \prod_{n=1}^{\infty} \frac{\alpha_n - z}{1 - \overline{\alpha}_n z} \frac{|\alpha_n|}{\alpha_n}$$

converges uniformly for $|z| \leq r < 1$, and defines a holomorphic function on the unit disc having precisely the zeros α_n and no other zeros. Show that $|f(z)| \leq 1$.

3.* Show that $\displaystyle\sum \frac{z^n}{(n!)^\alpha}$ is an entire function of order $1/\alpha$.

4.* Let $F(z) = \sum_{n=0}^{\infty} a_n z^n$ be an entire function of finite order. Then the growth order of F is intimately linked with the growth of the coefficients a_n as $n \to \infty$. In fact:

(a) Suppose $|F(z)| \leq A e^{a|z|^\rho}$. Then

$$\limsup_{n \to \infty} |a_n|^{1/n} n^{1/\rho} < \infty. \tag{8}$$

(b) Conversely, if (8) holds, then $|F(z)| \leq A_\epsilon e^{a_\epsilon |z|^{\rho+\epsilon}}$, for every $\epsilon > 0$.

[Hint: To prove (a), use Cauchy's inequality

$$|a_n| \leq \frac{A}{r^n} e^{ar^\rho},$$

and the fact that the function $u^{-n} e^{u^\rho}$, $0 < u < \rho$, attains its minimum value $e^{n/\rho}(\rho/n)^{n/\rho}$ at $u = n^{1/\rho}/\rho^{1/\rho}$. Then, choose r in terms of n to achieve this minimum.

To establish (b), note that for $|z| = r$,

$$|F(z)| \leq \sum \frac{c^n r^n}{n^{n/\rho}} \leq \sum \frac{c^n r^n}{(n!)^{1/\rho}}$$

for some constant c, since $n^n \geq n!$. This yields a reduction to Problem 3.]

6 The Gamma and Zeta Functions

It is no exaggeration to say that the gamma and zeta functions are among the most important nonelementary functions in mathematics. The gamma function Γ is ubiquitous in nature. It arises in a host of calculations and is featured in a large number of identities that occur in analysis. Part of the explanation for this probably lies in the basic structural property of the gamma function, which essentially characterizes it: $1/\Gamma(s)$ is the (simplest) entire function[1] which has zeros at exactly $s = 0, -1, -2, \ldots$.

The zeta function ζ (whose study, like that of the gamma function, was initiated by Euler) plays a fundamental role in the analytic theory of numbers. Its intimate connection with prime numbers comes about via the identity for $\zeta(s)$:

$$\prod_p \frac{1}{1-p^{-s}} = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

where the product is over all primes. The behavior of $\zeta(s)$ for real $s > 1$, with s tending to 1, was used by Euler to prove that $\sum_p 1/p$ diverges, and a similar reasoning for L-functions is at the starting point of the proof of Dirichlet's theorem on primes in arithmetic progression, as we saw in Book I.

While there is no difficulty in seeing that $\zeta(s)$ is well-defined (and analytic) when $\mathrm{Re}(s) > 1$, it was Riemann who realized that the further study of primes was bound up with the analytic (in fact, meromorphic) continuation of ζ into the rest of the complex plane. Beyond this, we also consider its remarkable functional equation, which reveals a symmetry about the line $\mathrm{Re}(s) = 1/2$, and whose proof is based on a corresponding identity for the theta function. We also make a more detailed study of the growth of $\zeta(s)$ near the line $\mathrm{Re}(s) = 1$, which will be required in the proof of the prime number theorem given in the next chapter.

---

[1]In keeping with the standard notation of the subject, we denote by s (instead of z) the argument of the functions Γ and ζ.

1 The gamma function

For $s > 0$, the **gamma function** is defined by

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt. \tag{1}$$

The integral converges for each positive s because near $t = 0$ the function t^{s-1} is integrable, and for t large the convergence is guaranteed by the exponential decay of the integrand. These observations allow us to extend the domain of definition of Γ as follows.

Proposition 1.1 *The gamma function extends to an analytic function in the half-plane* $\mathrm{Re}(s) > 0$*, and is still given there by the integral formula* (1)*.*

Proof. It suffices to show that the integral defines a holomorphic function in every strip

$$S_{\delta,M} = \{\delta < \mathrm{Re}(s) < M\},$$

where $0 < \delta < M < \infty$. Note that if σ denotes the real part of s, then $|e^{-t}t^{s-1}| = e^{-t}t^{\sigma-1}$, so that the integral

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt,$$

which is defined by the limit $\lim_{\epsilon\to 0} \int_\epsilon^{1/\epsilon} e^{-t}t^{s-1}\, dt$, converges for each $s \in S_{\delta,M}$. For $\epsilon > 0$, let

$$F_\epsilon(s) = \int_\epsilon^{1/\epsilon} e^{-t}t^{s-1}\, dt.$$

By Theorem 5.4 in Chapter 2, the function F_ϵ is holomorphic in the strip $S_{\delta,M}$. By Theorem 5.2, also of Chapter 2, it suffices to show that F_ϵ converges uniformly to Γ on the strip $S_{\delta,M}$. To see this, we first observe that

$$|\Gamma(s) - F_\epsilon(s)| \le \int_0^\epsilon e^{-t}t^{\sigma-1}\, dt + \int_{1/\epsilon}^\infty e^{-t}t^{\sigma-1}\, dt.$$

The first integral converges uniformly to 0, as ϵ tends to 0 since it can be easily estimated by ϵ^δ/δ whenever $0 < \epsilon < 1$. The second integral converges uniformly to 0 as well, since

$$\left|\int_{1/\epsilon}^\infty e^{-t}t^{\sigma-1}\, dt\right| \le \int_{1/\epsilon}^\infty e^{-t}t^{M-1} dt \le C\int_{1/\epsilon}^\infty e^{-t/2}\, dt \to 0,$$

and the proof is complete.

1.1 Analytic continuation

Despite the fact that the integral defining Γ is not absolutely convergent for other values of s, we can go further and prove that there exists a meromorphic function defined on all of $\mathbb{C}$ that equals Γ in the half-plane $\mathrm{Re}(s) > 0$. In the same sense as in Chapter 2, we say that this function is the analytic continuation[2] of Γ, and we therefore continue to denote it by Γ.

To prove the asserted analytic extension to a meromorphic function, we need a lemma, which incidentally exhibits an important property of Γ.

Lemma 1.2 *If* $\mathrm{Re}(s) > 0$*, then*

$$\Gamma(s+1) = s\Gamma(s). \tag{2}$$

As a consequence $\Gamma(n+1) = n!$ *for* $n = 0, 1, 2, \ldots$.

Proof. Integrating by parts in the finite integrals gives

$$\int_\epsilon^{1/\epsilon} \frac{d}{dt}(e^{-t}t^s)\,dt = -\int_\epsilon^{1/\epsilon} e^{-t}t^s\,dt + s\int_\epsilon^{1/\epsilon} e^{-t}t^{s-1}\,dt,$$

and the desired formula (2) follows by letting ϵ tend to 0, and noting that the left-hand side vanishes because $e^{-t}t^s \to 0$ as t tends to 0 or ∞. Now it suffices to check that

$$\Gamma(1) = \int_0^\infty e^{-t}\,dt = \left[-e^{-t}\right]_0^\infty = 1\,,$$

and to apply (2) successively to find that $\Gamma(n+1) = n!$.

Formula (2) in the lemma is all we need to give a proof of the following theorem.

Theorem 1.3 *The function* $\Gamma(s)$ *initially defined for* $\mathrm{Re}(s) > 0$ *has an analytic continuation to a meromorphic function on* $\mathbb{C}$ *whose only singularities are simple poles at the negative integers* $s = 0, -1, \ldots$. *The residue of* Γ *at* $s = -n$ *is* $(-1)^n/n!$.

[2]Uniqueness of the analytic continuation is guaranteed since the complement of the poles of a meromorphic function forms a connected set.

Proof. It suffices to extend Γ to each half-plane $\mathrm{Re}(s) > -m$, where $m \geq 1$ is an integer. For $\mathrm{Re}(s) > -1$, we define

$$F_1(s) = \frac{\Gamma(s+1)}{s}.$$

Since $\Gamma(s+1)$ is holomorphic in $\mathrm{Re}(s) > -1$, we see that F_1 is meromorphic in that half-plane, with the only possible singularity a simple pole at $s = 0$. The fact that $\Gamma(1) = 1$ shows that F_1 does in fact have a simple pole at $s = 0$ with residue 1. Moreover, if $\mathrm{Re}(s) > 0$, then

$$F_1(s) = \frac{\Gamma(s+1)}{s} = \Gamma(s)$$

by the previous lemma. So F_1 extends Γ to a meromorphic function on the half-plane $\mathrm{Re}(s) > -1$. We can now continue in this fashion by defining a meromorphic F_m for $\mathrm{Re}(s) > -m$ that agrees with Γ on $\mathrm{Re}(s) > 0$. For $\mathrm{Re}(s) > -m$, where m is an integer ≥ 1, define

$$F_m(s) = \frac{\Gamma(s+m)}{(s+m-1)(s+m-2)\cdots s}.$$

The function F_m is meromorphic in $\mathrm{Re}(s) > -m$ and has simple poles at $s = 0, -1, -2, \ldots, -m+1$ with residues

$$\begin{aligned}
\mathrm{res}_{s=-n} F_m(s) &= \frac{\Gamma(-n+m)}{(m-1-n)!(-1)(-2)\cdots(-n)} \\
&= \frac{(m-n-1)!}{(m-1-n)!(-1)(-2)\cdots(-n)} \\
&= \frac{(-1)^n}{n!}.
\end{aligned}$$

Successive applications of the lemma show that $F_m(s) = \Gamma(s)$ for $\mathrm{Re}(s) > 0$. By uniqueness, this also means that $F_m = F_k$ for $1 \leq k \leq m$ on the domain of definition of F_k. Therefore, we have obtained the desired continuation of Γ.

Remark. We have already proved that $\Gamma(s+1) = s\Gamma(s)$ whenever $\mathrm{Re}(s) > 0$. In fact, by analytic continuation, this formula remains true whenever $s \neq 0, -1, -2, \ldots$, that is, whenever s is not a pole of Γ. This is because both sides of the formula are holomorphic in the complement of the poles of Γ and are equal when $\mathrm{Re}(s) > 0$. Actually, one can go further, and note that if s is a negative integer $s = -n$ with $n \geq 1$, then both sides of the formula are infinite and moreover

$$\mathrm{res}_{s=-n}\Gamma(s+1) = -n\ \mathrm{res}_{s=-n}\Gamma(s).$$

Finally, note that when $s = 0$ we have $\Gamma(1) = \lim_{s\to 0} s\Gamma(s)$.

An alternate proof of Theorem 1.3, which is interesting in its own right and whose ideas recur later, is obtained by splitting the integral for $\Gamma(s)$ defined on $\text{Re}(s) > 0$ as follows:

$$\Gamma(s) = \int_0^1 e^{-t}t^{s-1}\,dt + \int_1^\infty e^{-t}t^{s-1}\,dt.$$

The integral on the far right defines an entire function; also expanding e^{-t} in a power series and integrating term by term gives

$$\int_0^1 e^{-t}t^{s-1}\,dt = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n+s)}.$$

Therefore

$$\Gamma(s) = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n+s)} + \int_1^\infty e^{-t}t^{s-1}\,dt \quad \text{for } \text{Re}(s) > 0. \tag{3}$$

Finally, the series defines a meromorphic function on $\mathbb{C}$ with poles at the negative integers and residue $(-1)^n/n!$ at $s = -n$. To prove this, we argue as follows. For a fixed $R > 0$ we may split the sum in two parts

$$\sum_{n=0}^\infty \frac{(-1)^n}{n!(n+s)} = \sum_{n=0}^N \frac{(-1)^n}{n!(n+s)} + \sum_{n=N+1}^\infty \frac{(-1)^n}{n!(n+s)},$$

where N is an integer chosen so that $N > 2R$. The first sum, which is finite, defines a meromorphic function in the disc $|s| < R$ with poles at the desired points and the correct residues. The second sum converges uniformly in that disc, hence defines a holomorphic function there, since $n > N > 2R$ and $|n + s| \geq R$ imply

$$\left|\frac{(-1)^n}{n!(n+s)}\right| \leq \frac{1}{n!R}.$$

Since R was arbitrary, we conclude that the series in (3) has the desired properties.

In particular, the relation (3) now holds on all of $\mathbb{C}$.

1.2 Further properties of Γ

The following identity reveals the symmetry of Γ about the line $\text{Re}(s) = 1/2$.

Theorem 1.4 *For all $s \in \mathbb{C}$,*

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}. \tag{4}$$

Observe that $\Gamma(1-s)$ has simple poles at the positive integers $s = 1, 2, 3, \ldots$, so that $\Gamma(s)\Gamma(1-s)$ is a meromorphic function on $\mathbb{C}$ with simple poles at *all* the integers, a property also shared by $\pi/\sin \pi s$.

To prove the identity, it suffices to do so for $0 < s < 1$ since it then holds on all of $\mathbb{C}$ by analytic continuation.

Lemma 1.5 *For $0 < a < 1$, $\displaystyle\int_0^\infty \frac{v^{a-1}}{1+v}\, dv = \frac{\pi}{\sin \pi a}$.*

Proof. We observe first that

$$\int_0^\infty \frac{v^{a-1}}{1+v}\, dv = \int_{-\infty}^\infty \frac{e^{ax}}{1+e^x}\, dx\,,$$

which follows by making the change of variables $v = e^x$. However, using contour integration, we saw in Example 2 of Section 2.1 in Chapter 3, that the second integral equals $\pi/\sin \pi a$, as desired.

To establish the theorem, we first note that for $0 < s < 1$ we may write

$$\Gamma(1-s) = \int_0^\infty e^{-u} u^{-s}\, du = t \int_0^\infty e^{-vt}(vt)^{-s}\, dv\,,$$

where for $t > 0$ we made the change of variables $vt = u$. This trick then gives

$$\begin{aligned}
\Gamma(1-s)\Gamma(s) &= \int_0^\infty e^{-t} t^{s-1} \Gamma(1-s)\, dt \\
&= \int_0^\infty e^{-t} t^{s-1} \left(t \int_0^\infty e^{-vt}(vt)^{-s} dv \right) dt \\
&= \int_0^\infty \int_0^\infty e^{-t[1+v]} v^{-s}\, dv dt \\
&= \int_0^\infty \frac{v^{-s}}{1+v}\, dv \\
&= \frac{\pi}{\sin \pi(1-s)} \\
&= \frac{\pi}{\sin \pi s},
\end{aligned}$$

and the theorem is proved.

In particular, by putting $s = 1/2$, and noting that $\Gamma(s) > 0$ whenever $s > 0$, we find that

$$\Gamma(1/2) = \sqrt{\pi}.$$

We continue our study of the gamma function by considering its reciprocal, which turns out to be an entire function with remarkably simple properties.

Theorem 1.6 *The function* Γ *has the following properties:*

(i) $1/\Gamma(s)$ *is an entire function of* s *with simple zeros at* $s=0,-1,-2,\ldots$ *and it vanishes nowhere else.*

(ii) $1/\Gamma(s)$ *has growth*

$$\left|\frac{1}{\Gamma(s)}\right| \leq c_1 e^{c_2 |s| \log |s|}.$$

Therefore, $1/\Gamma$ *is of order* 1 *in the sense that for every* $\epsilon > 0$*, there exists a bound* $c(\epsilon)$ *so that*

$$\left|\frac{1}{\Gamma(s)}\right| \leq c(\epsilon) e^{c_2 |s|^{1+\epsilon}}.$$

Proof. By the theorem we may write

$$\frac{1}{\Gamma(s)} = \Gamma(1-s)\frac{\sin \pi s}{\pi}, \tag{5}$$

so the simple poles of $\Gamma(1-s)$, which are at $s = 1, 2, 3, \ldots$ are cancelled by the simple zeros of $\sin \pi s$, and therefore $1/\Gamma$ is entire with simple zeros at $s = 0, -1, -2, -3, \ldots$.

To prove the estimate, we begin by showing that

$$\int_1^\infty e^{-t} t^\sigma \, dt \leq e^{(\sigma+1)\log(\sigma+1)}$$

whenever $\sigma = \mathrm{Re}(s)$ is positive. Choose n so that $\sigma \leq n \leq \sigma + 1$. Then

$$\begin{aligned}
\int_1^\infty e^{-t} t^\sigma \, dt &\leq \int_0^\infty e^{-t} t^n \, dt \\
&= n! \\
&\leq n^n \\
&= e^{n \log n} \\
&\leq e^{(\sigma+1)\log(\sigma+1)}.
\end{aligned}$$

Since the relation (3) holds on all of $\mathbb{C}$, we see from (5) that

$$\frac{1}{\Gamma(s)} = \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+1-s)}\right) \frac{\sin \pi s}{\pi} + \left(\int_1^\infty e^{-t} t^{-s}\, dt\right) \frac{\sin \pi s}{\pi}.$$

However, from our previous observation,

$$\left|\int_1^\infty e^{-t} t^{-s}\, dt\right| \leq \int_1^\infty e^{-t} t^{|\sigma|}\, dt \leq e^{(|\sigma|+1)\log(|\sigma|+1)},$$

and because $|\sin \pi s| \leq e^{\pi|s|}$ (by Euler's formula for the sine function) we find that the second term in the formula for $1/\Gamma(s)$ is dominated by $ce^{(|s|+1)\log(|s|+1)}e^{\pi|s|}$, which is itself majorized by $c_1 e^{c_2|s|\log|s|}$. Next, we consider the term

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+1-s)} \frac{\sin \pi s}{\pi}.$$

There are two cases: $|\text{Im}(s)| > 1$ and $|\text{Im}(s)| \leq 1$. In the first case, this expression is dominated in absolute value by $ce^{\pi|s|}$. If $|\text{Im}(s)| \leq 1$, we choose k to be the integer so that $k - 1/2 \leq \text{Re}(s) < k + 1/2$. Then if $k \geq 1$,

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+1-s)} \frac{\sin \pi s}{\pi} = (-1)^{k-1} \frac{\sin \pi s}{(k-1)!(k-s)\pi} +$$

$$+ \sum_{n \neq k-1} (-1)^n \frac{\sin \pi s}{n!(n+1-s)\pi}.$$

Both terms on the right are bounded; the first because $\sin \pi s$ vanishes at $s = k$, and the second because the sum is majorized by $c \sum 1/n!$.

When $k \leq 0$, then $\text{Re}(s) < 1/2$ by our supposition, and $\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+1-s)}$ is again bounded by $c \sum 1/n!$. This concludes the proof of the theorem.

The fact that $1/\Gamma$ satisfies the type of growth conditions discussed in Chapter 5 leads naturally to the product formula for the function $1/\Gamma$, which we treat next.

Theorem 1.7 *For all $s \in \mathbb{C}$,*

$$\frac{1}{\Gamma(s)} = e^{\gamma s} s \prod_{n=1}^{\infty} \left(1 + \frac{s}{n}\right) e^{-s/n}.$$

The real number γ, which is known as **Euler's constant**, is defined by

$$\gamma = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} - \log N.$$

The existence of the limit was already proved in Proposition 3.10, Chapter 8 of Book I, but we shall repeat the argument here for completeness. Observe that

$$\sum_{n=1}^{N} \frac{1}{n} - \log N = \sum_{n=1}^{N} \frac{1}{n} - \int_1^N \frac{1}{x}\,dx = \sum_{n=1}^{N-1} \int_n^{n+1} \left[\frac{1}{n} - \frac{1}{x}\right] dx + \frac{1}{N},$$

and by the mean value theorem applied to $f(x) = 1/x$ we have

$$\left|\frac{1}{n} - \frac{1}{x}\right| \le \frac{1}{n^2} \qquad \text{for all } n \le x \le n+1.$$

Hence

$$\sum_{n=1}^{\infty} \frac{1}{n} - \log N = \sum_{n=1}^{N-1} a_n + \frac{1}{N}$$

where $|a_n| \le 1/n^2$. Therefore $\sum a_n$ converges, which proves that the limit defining γ exists. We may now proceed with the proof of the factorization of $1/\Gamma$.

Proof. By the Hadamard factorization theorem and the fact that $1/\Gamma$ is entire, of growth order 1, and has simple zeros at $s = 0, -1, -2, \ldots$, we can expand $1/\Gamma$ in a Weierstrass product of the form

$$\frac{1}{\Gamma(s)} = e^{As+B} s \prod_{n=1}^{\infty} \left(1 + \frac{s}{n}\right) e^{-s/n}.$$

Here A and B are two constants that are to be determined. Remembering that $s\Gamma(s) \to 1$ as $s \to 0$, we find that $B = 0$ (or some integer multiple of $2\pi i$, which of course gives the same result). Putting $s = 1$, and using

the fact that $\Gamma(1) = 1$ yields

$$\begin{aligned} e^{-A} &= \prod_{n=1}^{\infty}\left(1+\frac{1}{n}\right)e^{-1/n} \\ &= \lim_{N\to\infty}\prod_{n=1}^{N}\left(1+\frac{1}{n}\right)e^{-1/n} \\ &= \lim_{N\to\infty} e^{\sum_{n=1}^{N}[\log(1+1/n)-1/n]} \\ &= \lim_{N\to\infty} e^{-\left(\sum_{n=1}^{N}1/n\right)+\log N+\log(1+1/N)} \\ &= e^{-\gamma}. \end{aligned}$$

Therefore $A = \gamma + 2\pi i k$ for some integer k. Since $\Gamma(s)$ is real whenever s is real, we must have $k = 0$, and the argument is complete.

Note that the proof shows that the function $1/\Gamma$ is essentially characterized (up to two normalizing constants) as the entire function that has:

(i) simple zeros at $s = 0, -1, -2, \ldots$ and vanishes nowhere else, and

(ii) order of growth ≤ 1.

Observe that $\sin \pi s$ has a similar characterization (except the zeros are now at *all* the integers). However, while $\sin \pi s$ has a stricter growth estimate of the form $\sin \pi s = O\left(e^{c|s|}\right)$, this estimate (without the logarithm in the exponent) does not hold for $1/\Gamma(s)$ as Exercise 12 demonstrates.

2 The zeta function

The Riemann **zeta function** is initially defined for real $s > 1$ by the convergent series

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s}.$$

As in the case of the gamma function, ζ can be continued into the complex plane. There are several proofs of this fact, and we present in the next section the one that relies on the functional equation of ζ.

2.1 Functional equation and analytic continuation

In parallel to the gamma function, we first provide a simple extension of ζ to a half-plane in $\mathbb{C}$.

Proposition 2.1 *The series defining $\zeta(s)$ converges for* $\mathrm{Re}(s) > 1$, *and the function ζ is holomorphic in this half-plane.*

Proof. If $s = \sigma + it$ where σ and t are real, then

$$|n^{-s}| = |e^{-s\log n}| = e^{-\sigma \log n} = n^{-\sigma}.$$

As a consequence, if $\sigma > 1 + \delta > 1$ the series defining ζ is uniformly bounded by $\sum_{n=1}^{\infty} 1/n^{1+\delta}$, which converges. Therefore, the series $\sum 1/n^s$ converges uniformly on every half-plane $\mathrm{Re}(s) > 1 + \delta > 1$, and therefore defines a holomorphic function in $\mathrm{Re}(s) > 1$.

The analytic continuation of ζ to a meromorphic function in $\mathbb{C}$ is more subtle than in the case of the gamma function. The proof we present here relates ζ to Γ and another important function.

Consider the **theta function**, already introduced in Chapter 4, which is defined for real $t > 0$ by

$$\vartheta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t}.$$

An application of the Poisson summation formula (Theorem 2.4 in Chapter 4) gave the functional equation satisfied by ϑ, namely

$$\vartheta(t) = t^{-1/2}\vartheta(1/t).$$

The growth and decay of ϑ we shall need are

$$\vartheta(t) \leq Ct^{-1/2} \qquad \text{as } t \to 0,$$

and

$$|\vartheta(t) - 1| \leq Ce^{-\pi t} \qquad \text{for some } C > 0, \text{ and all } t \geq 1.$$

The inequality for t tending to zero follows from the functional equation, while the behavior as t tends to infinity follows from the fact that

$$\sum_{n\geq 1} e^{-\pi n^2 t} \leq \sum_{n \geq 1} e^{-\pi n t} \leq Ce^{-\pi t}$$

for $t \geq 1$.

We are now in a position to prove an important relation among ζ, Γ and ϑ.

Theorem 2.2 *If* $\mathrm{Re}(s) > 1$*, then*

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{2}\int_0^\infty u^{(s/2)-1}[\vartheta(u) - 1]\, du.$$

Proof. This and further arguments are based on the observation that

$$\int_0^\infty e^{-\pi n^2 u} u^{(s/2)-1}\, du = \pi^{-s/2}\Gamma(s/2)n^{-s}, \quad \text{if } n \geq 1. \tag{6}$$

Indeed, if we make the change of variables $u = t/\pi n^2$ in the integral, the left-hand side becomes

$$\left(\int_0^\infty e^{-t} t^{(s/2)-1}\, dt\right)(\pi n^2)^{-s/2},$$

which is precisely $\pi^{-s/2}\Gamma(s/2)n^{-s}$. Next, note that

$$\frac{\vartheta(u) - 1}{2} = \sum_{n=1}^\infty e^{-\pi n^2 u}.$$

The estimates for ϑ given before the statement of the theorem justify an interchange of the infinite sum with the integral, and thus

$$\begin{aligned}
\frac{1}{2}\int_0^\infty u^{(s/2)-1}[\vartheta(u) - 1]\, du &= \sum_{n=1}^\infty \int_0^\infty u^{(s/2)-1} e^{-\pi n^2 u}\, du \\
&= \pi^{-s/2}\Gamma(s/2)\sum_{n=1}^\infty n^{-s} \\
&= \pi^{-s/2}\Gamma(s/2)\zeta(s),
\end{aligned}$$

as was to be shown.

In view of this, we now consider the modification of the ζ function called the **xi function**, which makes the former appear more symmetric. It is defined for $\mathrm{Re}(s) > 1$ by

$$\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s). \tag{7}$$

Theorem 2.3 *The function* ξ *is holomorphic for* $\mathrm{Re}(s) > 1$ *and has an analytic continuation to all of* $\mathbb{C}$ *as a meromorphic function with simple poles at* $s = 0$ *and* $s = 1$*. Moreover,*

$$\xi(s) = \xi(1 - s) \quad \textit{for all } s \in \mathbb{C}.$$

Proof. The idea of the proof is to use the functional equation for ϑ, namely

$$\sum_{n=-\infty}^{\infty} e^{-\pi n^2 u} = u^{-1/2} \sum_{n=-\infty}^{\infty} e^{-\pi n^2/u}, \qquad u > 0.$$

We then could multiply both sides by $u^{(s/2)-1}$ and try to integrate in u. Disregarding the terms corresponding to $n = 0$ (which produce infinities in both sums), we would get the desired equality once we invoked formula (6), and the parallel formula obtained by making the change of variables $u \mapsto 1/u$. The actual proof requires a little more work and goes as follows.

Let $\psi(u) = [\vartheta(u) - 1]/2$. The functional equation for the theta function, namely $\vartheta(u) = u^{-1/2}\vartheta(1/u)$, implies

$$\psi(u) = u^{-1/2}\psi(1/u) + \frac{1}{2u^{1/2}} - \frac{1}{2}.$$

Now, by Theorem 2.2 for $\mathrm{Re}(s) > 1$, we have

$$\begin{aligned}
\pi^{-s/2}\Gamma(s/2)\zeta(s) &= \int_0^\infty u^{(s/2)-1}\psi(u)\,du \\
&= \int_0^1 u^{(s/2)-1}\psi(u)\,du + \int_1^\infty u^{(s/2)-1}\psi(u)\,du \\
&= \int_0^1 u^{(s/2)-1}\left[u^{-1/2}\psi(1/u) + \frac{1}{2u^{1/2}} - \frac{1}{2}\right] du + \\
&\quad + \int_1^\infty u^{(s/2)-1}\psi(u)\,du \\
&= \frac{1}{s-1} - \frac{1}{s} + \int_1^\infty \left(u^{(-s/2)-1/2} + u^{(s/2)-1}\right)\psi(u)\,du
\end{aligned}$$

whenever $\mathrm{Re}(s) > 1$. Therefore

$$\xi(s) = \frac{1}{s-1} - \frac{1}{s} + \int_1^\infty \left(u^{(-s/2)-1/2} + u^{(s/2)-1}\right)\psi(u)\,du.$$

Since the function ψ has exponential decay at infinity, the integral above defines an entire function, and we conclude that ξ has an analytic continuation to all of $\mathbb{C}$ with simple poles at $s = 0$ and $s = 1$. Moreover, it is

immediate that the integral remains unchanged if we replace s by $1 - s$, and the same is true for the sum of the two terms $1/(s-1) - 1/s$. We conclude that $\xi(s) = \xi(1-s)$ as was to be shown.

From the identity we have proved for ξ we obtain the desired result for the zeta function: its analytic continuation and its functional equation.

Theorem 2.4 *The zeta function has a meromorphic continuation into the entire complex plane, whose only singularity is a simple pole at $s = 1$.*

Proof. A look at (7) provides the meromorphic continuation of ζ, namely

$$\zeta(s) = \pi^{s/2} \frac{\xi(s)}{\Gamma(s/2)}.$$

Recall that $1/\Gamma(s/2)$ is entire with simple zeros at $0, -2, -4, \ldots$, so the simple pole of $\xi(s)$ at the origin is cancelled by the corresponding zero of $1/\Gamma(s/2)$. As a consequence, the only singularity of ζ is a simple pole at $s = 1$.

We shall now present a more elementary approach to the analytic continuation of the zeta function, which easily leads to its extension in the half-plane $\text{Re}(s) > 0$. This method will be useful in studying the growth properties of ζ near the line $\text{Re}(s) = 1$ (which will be needed in the next chapter). The idea behind it is to compare the sum $\sum_{n=1}^{\infty} n^{-s}$ with the integral $\int_1^{\infty} x^{-s}\, dx$.

Proposition 2.5 *There is a sequence of entire functions $\{\delta_n(s)\}_{n=1}^{\infty}$ that satisfy the estimate $|\delta_n(s)| \leq |s|/n^{\sigma+1}$, where $s = \sigma + it$, and such that*

$$\sum_{1 \leq n < N} \frac{1}{n^s} - \int_1^N \frac{dx}{x^s} = \sum_{1 \leq n < N} \delta_n(s), \tag{8}$$

whenever N is an integer > 1.

This proposition has the following consequence.

Corollary 2.6 *For* $\text{Re}(s) > 0$ *we have*

$$\zeta(s) - \frac{1}{s-1} = H(s),$$

where $H(s) = \sum_{n=1}^{\infty} \delta_n(s)$ is holomorphic in the half-plane $\text{Re}(s) > 0$.

To prove the proposition we compare $\sum_{1\leq n<N} n^{-s}$ with $\sum_{1\leq n<N}\int_n^{n+1} x^{-s}\,dx$, and set

$$\delta_n(s) = \int_n^{n+1}\left[\frac{1}{n^s}-\frac{1}{x^s}\right]dx. \tag{9}$$

The mean-value theorem applied to $f(x) = x^{-s}$ yields

$$\left|\frac{1}{n^s}-\frac{1}{x^s}\right| \leq \frac{|s|}{n^{\sigma+1}}, \qquad \text{whenever } n \leq x \leq n+1.$$

Therefore $|\delta_n(s)| \leq |s|/n^{\sigma+1}$, and since

$$\int_1^N \frac{dx}{x^s} = \sum_{1\leq n<N}\int_n^{n+1}\frac{dx}{x^s},$$

the proposition is proved.

Turning to the corollary, we assume first that $\mathrm{Re}(s) > 1$. We let N tend to infinity in formula (8) of the proposition, and observe that by the estimate $|\delta_n(s)| \leq |s|/n^{\sigma+1}$ we have the uniform convergence of the series $\sum \delta_n(s)$ (in any half-plane $\mathrm{Re}(s) \geq \delta$ when $\delta > 0$). Since $\mathrm{Re}(s) > 1$, the series $\sum n^{-s}$ converges to $\zeta(s)$, and this proves the assertion when $\mathrm{Re}(s) > 1$. The uniform convergence also shows that $\sum \delta_n(s)$ is holomorphic when $\mathrm{Re}(s) > 0$, and thus shows that $\zeta(s)$ is extendable to that half-plane, and that the identity continues to hold there.

Remark. The idea described above can be developed step by step to yield the continuation of ζ into the entire complex plane, as shown in Problems 2 and 3. Another argument giving the full analytic continuation of ζ is outlined in Exercises 15 and 16.

As an application of the proposition we can show that the growth of $\zeta(s)$ near the line $\mathrm{Re}(s) = 1$ is "mild." Recall that when $\mathrm{Re}(s) > 1$, we have $|\zeta(s)| \leq \sum_{n=1}^{\infty} n^{-\sigma}$, and so $\zeta(s)$ is bounded in any half-plane $\mathrm{Re}(s) \geq 1+\delta$, with $\delta > 0$. We shall see that on the line $\mathrm{Re}(s) = 1$, $|\zeta(s)|$ is majorized by $|t|^{\epsilon}$, for every $\epsilon > 0$, and that the growth near the line is not much worse. The estimates below are not optimal. In fact, they are rather crude but suffice for what is needed later on.

Proposition 2.7 *Suppose $s = \sigma + it$ with $\sigma, t \in \mathbb{R}$. Then for each σ_0, $0 \leq \sigma_0 \leq 1$, and every $\epsilon > 0$, there exists a constant c_ϵ so that*

(i) $|\zeta(s)| \leq c_\epsilon |t|^{1-\sigma_0+\epsilon}$, *if $\sigma_0 \leq \sigma$ and $|t| \geq 1$.*

(ii) $|\zeta'(s)| \le c_\epsilon |t|^\epsilon$, *if* $1 \le \sigma$ *and* $|t| \ge 1$.

In particular, the proposition implies that $\zeta(1+it) = O(|t|^\epsilon)$ as $|t|$ tends to infinity,[3] and the same estimate also holds for ζ'. For the proof, we use Corollary 2.6. Recall the estimate $|\delta_n(s)| \le |s|/n^{\sigma+1}$. We also have the estimate $|\delta_n(s)| \le 2/n^\sigma$, which follows from the expression for $\delta_n(s)$ given by (9) and the fact that $|n^{-s}| = n^{-\sigma}$ and $|x^{-s}| \le n^{-\sigma}$ if $x \ge n$. We then combine these two estimates for $|\delta_n(s)|$ via the observation that $A = A^\delta A^{1-\delta}$, to obtain the bound

$$|\delta_n(s)| \le \left(\frac{|s|}{n^{\sigma_0+1}}\right)^\delta \left(\frac{2}{n^{\sigma_0}}\right)^{1-\delta} \le \frac{2|s|^\delta}{n^{\sigma_0+\delta}},$$

as long as $\delta \ge 0$. Now choose $\delta = 1 - \sigma_0 + \epsilon$ and apply the identity in Corollary 2.6. Then, with $\sigma = \mathrm{Re}(s) \ge \sigma_0$, we find

$$|\zeta(s)| \le \left|\frac{1}{s-1}\right| + 2|s|^{1-\sigma_0+\epsilon} \sum_{n=1}^\infty \frac{1}{n^{1+\epsilon}},$$

and conclusion (i) is proved. The second conclusion is actually a consequence of the first by a slight modification of Exercise 8 in Chapter 2. For completeness we sketch the argument. By the Cauchy integral formula,

$$\zeta'(s) = \frac{1}{2\pi r} \int_0^{2\pi} \zeta(s + re^{i\theta}) e^{i\theta}\, d\theta,$$

where the integration is taken over a circle of radius r centered at the point s. Now choose $r = \epsilon$ and observe that this circle lies in the half-plane $\mathrm{Re}(s) \ge 1 - \epsilon$, and so (ii) follows as a consequence of (i) on replacing 2ϵ by ϵ.

3 Exercises

1. Prove that

$$\Gamma(s) = \lim_{n\to\infty} \frac{n^s n!}{s(s+1)\cdots(s+n)}$$

whenever $s \ne 0, -1, -2, \ldots$.

[Hint: Use the product formula for $1/\Gamma$, and the definition of the Euler constant γ.]

2. Prove that

$$\prod_{n=1}^\infty \frac{n(n+a+b)}{(n+a)(n+b)} = \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+1)}$$

[3] The reader should recall the O notation which was introduced at the end of Chapter 1.

whenever a and b are positive. Using the product formula for $\sin \pi s$, give another proof that $\Gamma(s)\Gamma(1-s) = \pi / \sin \pi s$.

3. Show that Wallis's product formula can be written as

$$\sqrt{\frac{\pi}{2}} = \lim_{n\to\infty} \frac{2^{2n}(n!)^2}{(2n+1)!}(2n+1)^{1/2}.$$

As a result, prove the following identity:

$$\Gamma(s)\Gamma(s+1/2) = \sqrt{\pi}2^{1-2s}\Gamma(2s).$$

4. Prove that if we take

$$f(z) = \frac{1}{(1-z)^\alpha}, \qquad \text{for } |z| < 1$$

(defined in terms of the principal branch of the logarithm), where α is a fixed complex number, then

$$f(z) = \sum_{n=0}^{\infty} a_n(\alpha) z^n$$

with

$$a_n(\alpha) \sim \frac{1}{\Gamma(\alpha)} n^{\alpha-1} \qquad \text{as } n \to \infty.$$

5. Use the fact that $\Gamma(s)\Gamma(1-s) = \pi / \sin \pi s$ to prove that

$$|\Gamma(1/2+it)| = \sqrt{\frac{2\pi}{e^{\pi t} + e^{-\pi t}}}, \qquad \text{whenever } t \in \mathbb{R}.$$

6. Show that

$$1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1} - \frac{1}{2}\log n \to \frac{\gamma}{2} + \log 2,$$

where γ is Euler's constant.

7. The **Beta function** is defined for $\mathrm{Re}(\alpha) > 0$ and $\mathrm{Re}(\beta) > 0$ by

$$B(\alpha, \beta) = \int_0^1 (1-t)^{\alpha-1} t^{\beta-1}\, dt.$$

(a) Prove that $B(\alpha, \beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

(b) Show that $B(\alpha,\beta)=\displaystyle\int_0^\infty \frac{u^{\alpha-1}}{(1+u)^{\alpha+\beta}}\,du.$

[Hint: For part (a), note that

$$\Gamma(\alpha)\Gamma(\beta)=\int_0^\infty\int_0^\infty t^{\alpha-1}s^{\beta-1}e^{-t-s}\,dtds,$$

and make the change of variables $s=ur$, $t=u(1-r)$.]

8. The Bessel functions arise in the study of spherical symmetries and the Fourier transform. See Chapter 6 in Book I. Prove that the following power series identity holds for Bessel functions of real order $\nu>-1/2$:

$$J_\nu(x)=\frac{(x/2)^\nu}{\Gamma(\nu+1/2)\sqrt{\pi}}\int_{-1}^1 e^{ixt}(1-t^2)^{\nu-(1/2)}\,dt=\left(\frac{x}{2}\right)^\nu\sum_{m=0}^\infty\frac{(-1)^m\left(\frac{x^2}{4}\right)^m}{m!\Gamma(\nu+m+1)}$$

whenever $x>0$. In particular, the Bessel function J_ν satisfies the ordinary differential equation

$$\frac{d^2J_\nu}{dx^2}+\frac{1}{x}\frac{dJ_\nu}{dx}+\left(1-\frac{\nu^2}{x^2}\right)J_\nu=0.$$

[Hint: Expand the exponential e^{ixt} in a power series, and express the remaining integrals in terms of the gamma function, using Exercise 7.]

9. The hypergeometric series $F(\alpha,\beta,\gamma;z)$ was defined in Exercise 16 of Chapter 1. Show that

$$F(\alpha,\beta,\gamma;z)=\frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)}\int_0^1 t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-zt)^{-\alpha}\,dt.$$

Here $\alpha>0,\beta>0,\gamma>\beta$, and $|z|<1$.

Show as a result that the hypergeometric function, initially defined by a power series convergent in the unit disc, can be continued analytically to the complex plane slit along the half-line $[1,\infty)$.

Note that

$$\begin{aligned}\log(1-z)&=-zF(1,1,2;z),\\ e^z&=\lim_{\beta\to\infty}F(1,\beta,1;z/\beta),\\ (1-z)^{-\alpha}&=F(\alpha,1,1;z).\end{aligned}$$

[Hint: To prove the integral identity, expand $(1-zt)^{-\alpha}$ as a power series.]

10. An integral of the form

$$F(z)=\int_0^\infty f(t)t^{z-1}\,dt$$

is called a **Mellin transform**, and we shall write $\mathcal{M}(f)(z) = F(z)$. For example, the gamma function is the Mellin transform of the function e^{-t}.

(a) Prove that

$$\mathcal{M}(\cos)(z) = \int_0^\infty \cos(t)t^{z-1}\,dt = \Gamma(z)\cos\left(\pi\frac{z}{2}\right) \qquad \text{for } 0 < \text{Re}(z) < 1,$$

and

$$\mathcal{M}(\sin)(z) = \int_0^\infty \sin(t)t^{z-1}\,dt = \Gamma(z)\sin\left(\pi\frac{z}{2}\right) \qquad \text{for } 0 < \text{Re}(z) < 1.$$

(b) Show that the second of the above identities is valid in the larger strip $-1 < \text{Re}(z) < 1$, and that as a consequence, one has

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2} \qquad \text{and} \qquad \int_0^\infty \frac{\sin x}{x^{3/2}}\,dx = \sqrt{2\pi}.$$

This generalizes the calculation in Exercise 2 of Chapter 2.

[Hint: For the first part, consider the integral of the function $f(w) = e^{-w}w^{z-1}$ around the contour illustrated in Figure 1. Use analytic continuation to prove the second part.]

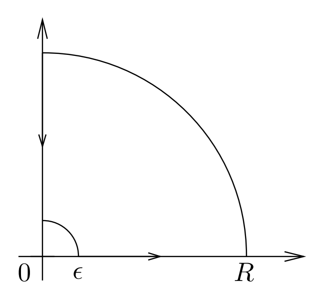


Figure 1. The contour in Exercise 10

11. Let $f(z) = e^{az}e^{-e^z}$ where $a > 0$. Observe that in the strip $\{x + iy : |y| < \pi\}$ the function $f(x + iy)$ is exponentially decreasing as $|x|$ tends to infinity. Prove that

$$\hat{f}(\xi) = \Gamma(a + i\xi), \qquad \text{for all } \xi \in \mathbb{R}.$$

12. This exercise gives two simple observations about $1/\Gamma$.

(a) Show that $1/|\Gamma(s)|$ is not $O(e^{c|s|})$ for any $c > 0$. [Hint: If $s = -k - 1/2$, where k is a positive integer, then $|1/\Gamma(s)| \geq k!/\pi$.]

(b) Show that there is no entire function $F(s)$ with $F(s) = O(e^{c|s|})$ that has simple zeros at $s = 0, -1, -2, \ldots, -n, \ldots$, and that vanishes nowhere else.

13. Prove that

$$\frac{d^2 \log \Gamma(s)}{ds^2} = \sum_{n=0}^{\infty} \frac{1}{(s+n)^2}$$

whenever s is a positive number. Show that if the left-hand side is interpreted as $(\Gamma'/\Gamma)'$, then the above formula also holds for all complex numbers s with $s \neq 0, -1, -2, \ldots$.

14. This exercise gives an asymptotic formula for $\log n!$. A more refined asymptotic formula for $\Gamma(s)$ as $s \to \infty$ (Stirling's formula) is given in Appendix A.

(a) Show that

$$\frac{d}{dx} \int_x^{x+1} \log \Gamma(t)\, dt = \log x, \qquad \text{for } x > 0,$$

and as a result

$$\int_x^{x+1} \log \Gamma(t)\, dt = x \log x - x + c.$$

(b) Show as a consequence that $\log \Gamma(n) \sim n \log n$ as $n \to \infty$. In fact, prove that $\log \Gamma(n) \sim n \log n + O(n)$ as $n \to \infty$. [Hint: Use the fact that $\Gamma(x)$ is monotonically increasing for all large x.]

15. Prove that for $\mathrm{Re}(s) > 1$,

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{x^{s-1}}{e^x - 1}\, dx.$$

[Hint: Write $1/(e^x - 1) = \sum_{n=1}^{\infty} e^{-nx}$.]

16. Use the previous exercise to give another proof that $\zeta(s)$ is continuable in the complex plane with only singularity a simple pole at $s = 1$.

[Hint: Write

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^1 \frac{x^{s-1}}{e^x - 1}\, dx + \frac{1}{\Gamma(s)} \int_1^{\infty} \frac{x^{s-1}}{e^x - 1}\, dx.$$

The second integral defines an entire function, while

$$\int_0^1 \frac{x^{s-1}}{e^x - 1}\,dx = \sum_{m=0}^{\infty} \frac{B_m}{m!(s+m-1)},$$

where B_m denotes the mth Bernoulli number defined by

$$\frac{x}{e^x - 1} = \sum_{m=0}^{\infty} \frac{B_m}{m!} x^m.$$

Then $B_0 = 1$, and since $z/(e^z - 1)$ is holomorphic for $|z| < 2\pi$, we must have $\limsup_{m\to\infty} |B_m/m!|^{1/m} = 1/2\pi$.]

17. Let f be an indefinitely differentiable function on $\mathbb{R}$ that has compact support, or more generally, let f belong to the Schwartz space.[4] Consider

$$I(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} f(x) x^{-1+s}\,dx.$$

(a) Observe that $I(s)$ is holomorphic for $\text{Re}(s) > 0$. Prove that I has an analytic continuation as an entire function in the complex plane.

(b) Prove that $I(0) = 0$, and more generally

$$I(-n) = (-1)^n f^{(n+1)}(0) \qquad \text{for all } n \geq 0.$$

[Hint: To prove the analytic continuation, as well as the formulas in the second part, integrate by parts to show that $I(s) = \frac{(-1)^k}{\Gamma(s+k)} \int_0^\infty f^{(k)}(x) x^{s+k-1}\,dx$.]

4 Problems

1. This problem provides further estimates for ζ and ζ' near $\text{Re}(s) = 1$.

(a) Use Proposition 2.5 and its corollary to prove

$$\zeta(s) = \sum_{1 \leq n < N} n^{-s} - \frac{N^{s-1}}{s-1} + \sum_{n \geq N} \delta_n(s)$$

for every integer $N \geq 2$, whenever $\text{Re}(s) > 0$.

(b) Show that $|\zeta(1+it)| = O(\log |t|)$, as $|t| \to \infty$ by using the previous result with N =greatest integer in $|t|$.

[4]The Schwartz space on $\mathbb{R}$ is denoted by $\mathcal{S}$ and consists of all indefinitely differentiable functions f, so that f and all its derivatives decay faster than any polynomials. In other words, $\sup_{x\in\mathbb{R}} |x|^m |f^{(\ell)}(x)| < \infty$ for all integers $m, \ell \geq 0$. This space appeared in the study of the Fourier transform in Book I.

(c) The second conclusion of Proposition 2.7 can be similarly refined.

(d) Show that if $t \neq 0$ and t is fixed, then the partial sums of the series $\sum_{n=1}^{\infty} 1/n^{1+it}$ are bounded, but the series does not converge.

2.* Prove that for $\text{Re}(s) > 0$

$$\zeta(s) = \frac{s}{s-1} - s \int_1^\infty \frac{\{x\}}{x^{s+1}}\, dx$$

where $\{x\}$ is the fractional part of x.

3.* If $Q(x) = \{x\} - 1/2$, then we can write the expression in the previous problem as

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s \int_1^\infty \frac{Q(x)}{x^{s+1}}\, dx.$$

Let us construct $Q_k(x)$ recursively so that

$$\int_0^1 Q_k(x)\, dx = 0, \qquad \frac{dQ_{k+1}}{dx} = Q_k(x), \qquad Q_0(x) = Q(x) \qquad \text{and} \qquad Q_k(x+1) = Q_k(x).$$

Then we can write

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s \int_1^\infty \left(\frac{d^k}{dx^k} Q_k(x) \right) x^{-s-1}\, dx\,,$$

and a k-fold integration by parts gives the analytic continuation for $\zeta(s)$ when $\text{Re}(s) > -k$.

4.* The functions Q_k in the previous problem are related to the Bernoulli polynomials $B_k(x)$ by the formula

$$Q_k(x) = \frac{B_{k+1}(x)}{(k+1)!} \qquad \text{for } 0 \leq x \leq 1.$$

Also, if k is a positive integer, then

$$2\zeta(2k) = (-1)^{k+1} \frac{(2\pi)^{2k}}{(2k)!} B_{2k},$$

where $B_k = B_k(0)$ are the Bernoulli numbers. For the definition of $B_k(x)$ and B_k see Chapter 3 in Book I.

7 The Zeta Function and Prime Number Theorem

> Bernhard Riemann, whose extraordinary intuitive powers we have already mentioned, has especially renovated our knowledge of the distribution of prime numbers, also one of the most mysterious questions in mathematics. He has taught us to deduce results in that line from considerations borrowed from the integral calculus: more precisely, from the study of a certain quantity, a function of a variable s which may assume not only real, but also imaginary values. He proved some important properties of that function, but enunciated two or three as important ones without giving the proof. At the death of Riemann, a note was found among his papers, saying "These properties of $\zeta(s)$ (the function in question) are deduced from an expression of it which, however, I did not succeed in simplifying enough to publish it."
>
> We still have not the slightest idea of what the expression could be. As to the properties he simply enunciated, some thirty years elapsed before I was able to prove all of them but one. The question concerning that last one remains unsolved as yet, though, by an immense labor pursued throughout this last half century, some highly interesting discoveries in that direction have been achieved. It seems more and more probable, but still not at all certain, that the "Riemann hypothesis" is true.
>
> *J. Hadamard,* 1945

Euler found, through his product formula for the zeta function, a deep connection between analytical methods and arithmetic properties of numbers, in particular primes. An easy consequence of Euler's formula is that the sum of the reciprocals of all primes, $\sum_p 1/p$, diverges, a result that quantifies the fact that there are infinitely many prime numbers. The natural problem then becomes that of understanding how these primes are distributed. With this in mind, we consider the

following function:

$$\pi(x) = \text{number of primes less than or equal to } x.$$

The erratic growth of the function $\pi(x)$ gives little hope of finding a simple formula for it. Instead, one is led to study the asymptotic behavior of $\pi(x)$ as x becomes large. About 60 years after Euler's discovery, Legendre and Gauss observed after numerous calculations that it was likely that

$$\pi(x) \sim \frac{x}{\log x} \qquad \text{as } x \to \infty. \tag{1}$$

(The asymptotic relation $f(x) \sim g(x)$ as $x \to \infty$ means that $f(x)/g(x) \to 1$ as $x \to \infty$.) Another 60 years later, shortly before Riemann's work, Tchebychev proved by elementary methods (and in particular, without the zeta function) the weaker result that

$$\pi(x) \approx \frac{x}{\log x} \qquad \text{as } x \to \infty. \tag{2}$$

Here, by definition, the symbol $\approx$ means that there are positive constants $A < B$ such that

$$A\,\frac{x}{\log x} \leq \pi(x) \leq B\,\frac{x}{\log x}$$

for all sufficiently large x.

In 1896, about 40 years after Tchebychev's result, Hadamard and de la Vallée Poussin gave a proof of the validity of the relation (1). Their result is known as the prime number theorem. The original proofs of this theorem, as well as the one we give below, use complex analysis. We should remark that since then other proofs have been found, some depending on complex analysis, and others more elementary in nature.

At the heart of the proof of the prime number theorem that we give below lies the fact that $\zeta(s)$ does not vanish on the line $\text{Re}(s) = 1$. In fact, it can be shown that these two propositions are equivalent.

1 Zeros of the zeta function

We have seen in Theorem 1.10, Chapter 8 in Book I, Euler's identity, which states that for $\text{Re}(s) > 1$ the zeta function can be expressed as an infinite product

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}.$$

For the sake of completeness we provide a proof of the above identity. The key observation is that $1/(1-p^{-s})$ can be written as a convergent (geometric) power series

$$1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\cdots+\frac{1}{p^{Ms}}+\cdots,$$

and taking formally the product of these series over all primes p, yields the desired result. A precise argument goes as follows.

Suppose M and N are positive integers with $M > N$. Observe now that, by the fundamental theorem of arithmetic,[1] any positive integer $n \leq N$ can be written uniquely as a product of primes, and that each prime that occurs in the product must be less than or equal to N and repeated less than M times. Therefore

$$\begin{aligned}\sum_{n=1}^{N}\frac{1}{n^s} &\leq \prod_{p\leq N}\left(1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\cdots+\frac{1}{p^{Ms}}\right)\\ &\leq \prod_{p\leq N}\left(\frac{1}{1-p^{-s}}\right)\\ &\leq \prod_{p}\left(\frac{1}{1-p^{-s}}\right).\end{aligned}$$

Letting N tend to infinity in the series now yields

$$\sum_{n=1}^{\infty}\frac{1}{n^s}\leq\prod_{p}\left(\frac{1}{1-p^{-s}}\right).$$

For the reverse inequality, we argue as follows. Again, by the fundamental theorem of arithmetic, we find that

$$\prod_{p\leq N}\left(1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\cdots+\frac{1}{p^{Ms}}\right)\leq\sum_{n=1}^{\infty}\frac{1}{n^s}.$$

Letting M tend to infinity gives

$$\prod_{p\leq N}\left(\frac{1}{1-p^{-s}}\right)\leq\sum_{n=1}^{\infty}\frac{1}{n^s}.$$

[1] A proof of this elementary (but essential) fact is given in the first section of Chapter 8 in Book I.

Hence

$$\prod_p \left(\frac{1}{1-p^{-s}}\right) \leq \sum_{n=1}^{\infty} \frac{1}{n^s},$$

and the proof of the product formula for ζ is complete.

From the product formula we see, by Proposition 3.1 in Chapter 5, that $\zeta(s)$ does not vanish when $\mathrm{Re}(s) > 1$.

To obtain further information about the location of the zeros of ζ, we use the functional equation that provided the analytic continuation of ζ. We may write the fundamental relation $\xi(s) = \xi(1-s)$ in the form

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-(1-s)/2}\Gamma((1-s)/2)\zeta(1-s),$$

and therefore

$$\zeta(s) = \pi^{s-1/2}\frac{\Gamma((1-s)/2)}{\Gamma(s/2)}\zeta(1-s).$$

Now observe that for $\mathrm{Re}(s) < 0$ the following are true:

(i) $\zeta(1-s)$ has no zeros because $\mathrm{Re}(1-s) > 1$.

(ii) $\Gamma((1-s)/2)$ is zero free.

(iii) $1/\Gamma(s/2)$ has zeros at $s = -2, -4, -6, \ldots$.

Therefore, the only zeros of ζ in $\mathrm{Re}(s) < 0$ are located at the negative even integers $-2, -4, -6, \ldots$.

This proves the following theorem.

Theorem 1.1 *The only zeros of* ζ *outside the strip* $0 \leq \mathrm{Re}(s) \leq 1$ *are at the negative even integers,* $-2, -4, -6, \ldots$.

The region that remains to be studied is called the **critical strip**, $0 \leq \mathrm{Re}(s) \leq 1$. A key fact in the proof of the prime number theorem is that ζ has no zeros on the line $\mathrm{Re}(s) = 1$. As a simple consequence of this fact and the functional equation, it follows that ζ has no zeros on the line $\mathrm{Re}(s) = 0$.

In the seminal paper where Riemann introduced the analytic continuation of the ζ function and proved its functional equation, he applied these insights to the theory of prime numbers, and wrote down "explicit" formulas for determining the distribution of primes. While he did not succeed in fully proving and exploiting his assertions, he did initiate many important new ideas. His analysis led him to believe the truth of what has since been called the **Riemann hypothesis**:

> *The zeros of $\zeta(s)$ in the critical strip lie on the line* $\text{Re}(s) = 1/2$.

He said about this: "It would certainly be desirable to have a rigorous demonstration of this proposition; nevertheless I have for the moment set this aside, after several quick but unsuccessful attempts, because it seemed unneeded for the immediate goal of my study." Although much of the theory and numerical results point to the validity of this hypothesis, a proof or a counter-example remains to be discovered. The Riemann hypothesis is today one of mathematics' most famous unresolved problems.

In particular, it is for this reason that the zeros of ζ located outside the critical strip are sometimes called the **trivial zeros** of the zeta function. See also Exercise 5 for an argument proving that ζ has no zeros on the real segment, $0 \leq \sigma \leq 1$, where $s = \sigma + it$.

In the rest of this section we shall restrict ourselves to proving the following theorem, together with related estimates on ζ, which we shall use in the proof of the prime number theorem.

Theorem 1.2 *The zeta function has no zeros on the line* $\text{Re}(s) = 1$.

Of course, since we know that ζ has a pole at $s = 1$, there are no zeros in a neighborhood of this point, but what we need is the deeper property that

$$\zeta(1 + it) \neq 0 \qquad \text{for all } t \in \mathbb{R}.$$

The next sequence of lemmas gathers the necessary ingredients for the proof of Theorem 1.2.

Lemma 1.3 *If* $\text{Re}(s) > 1$, *then*

$$\log \zeta(s) = \sum_{p,m} \frac{p^{-ms}}{m} = \sum_{n=1}^{\infty} c_n n^{-s}$$

for some $c_n \geq 0$.

Proof. Suppose first that $s > 1$. Taking the logarithm of the Euler product formula, and using the power series expansion for the logarithm

$$\log\left(\frac{1}{1-x}\right) = \sum_{m=1}^{\infty} \frac{x^m}{m},$$

which holds for $0 \leq x < 1$, we find that

$$\log \zeta(s) = \log \prod_p \frac{1}{1-p^{-s}} = \sum_p \log \left(\frac{1}{1-p^{-s}} \right) = \sum_{p,m} \frac{p^{-ms}}{m}.$$

Since the double sum converges absolutely, we need not specify the order of summation. See the Note at the end of this chapter. The formula then holds for all $\text{Re}(s) > 1$ by analytic continuation. Note that, by Theorem 6.2 in Chapter 3, $\log \zeta(s)$ is well defined in the simply connected half-plane $\text{Re}(s) > 1$, since ζ has no zeros there. Finally, it is clear that we have

$$\sum_{p,m} \frac{p^{-ms}}{m} = \sum_{n=1}^{\infty} c_n n^{-s},$$

where $c_n = 1/m$ if $n = p^m$ and $c_n = 0$ otherwise.

The proof of the theorem we shall give depends on a simple trick that is based on the following inequality.

Lemma 1.4 *If* $\theta \in \mathbb{R}$, *then* $3 + 4\cos\theta + \cos 2\theta \geq 0$.

This follows at once from the simple observation

$$3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2.$$

Corollary 1.5 *If* $\sigma > 1$ *and* t *is real, then*

$$\log |\zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it)| \geq 0.$$

Proof. Let $s = \sigma + it$ and note that

$$\text{Re}(n^{-s}) = \text{Re}(e^{-(\sigma+it)\log n}) = e^{-\sigma \log n} \cos(t \log n) = n^{-\sigma} \cos(t \log n).$$

Therefore,

$$\begin{aligned}
&\log |\zeta^3(\sigma)\zeta^4(\sigma + it)\zeta(\sigma + 2it)| \\
&= 3 \log |\zeta(\sigma)| + 4 \log |\zeta(\sigma + it)| + \log |\zeta(\sigma + 2it)| \\
&= 3\text{Re}[\log \zeta(\sigma)] + 4\text{Re}[\log \zeta(\sigma + it)] + \text{Re}[\log \zeta(\sigma + 2it)] \\
&= \sum c_n n^{-\sigma} (3 + 4\cos\theta_n + \cos 2\theta_n),
\end{aligned}$$

where $\theta_n = t \log n$. The positivity now follows from Lemma 1.4, and the fact that $c_n \geq 0$.

We can now finish the proof of our theorem.

Proof of Theorem 1.2. Suppose on the contrary that $\zeta(1+it_0)=0$ for some $t_0 \neq 0$. Since ζ is holomorphic at $1+it_0$, it must vanish at least to order 1 at this point, hence

$$|\zeta(\sigma+it_0)|^4 \leq C(\sigma-1)^4 \quad \text{as } \sigma \to 1,$$

for some constant $C>0$. Also, we know that $s=1$ is a simple pole for $\zeta(s)$, so that

$$|\zeta(\sigma)|^3 \leq C'(\sigma-1)^{-3} \quad \text{as } \sigma \to 1,$$

for some constant $C'>0$. Finally, since ζ is holomorphic at the points $\sigma+2it_0$, the quantity $|\zeta(\sigma+2it_0)|$ remains bounded as $\sigma \to 1$. Putting these facts together yields

$$|\zeta^3(\sigma)\zeta^4(\sigma+it)\zeta(\sigma+2it)| \to 0 \quad \text{as } \sigma \to 1\,,$$

which contradicts Corollary 1.5, since the logarithm of real numbers between 0 and 1 is negative. This concludes the proof that ζ is zero free on the real line $\text{Re}(s)=1$.

1.1 Estimates for $1/\zeta(s)$

The proof of the prime number theorem relies on detailed manipulations of the zeta function near the line $\text{Re}(s)=1$; the basic object involved is the logarithmic derivative $\zeta'(s)/\zeta(s)$. For this reason, besides the non-vanishing of ζ on the line, we need to know about the growth of ζ' and $1/\zeta$. The former was dealt with in Proposition 2.7 of Chapter 6; we now treat the latter.

The proposition that follows is actually a quantitative version of Theorem 1.2.

Proposition 1.6 *For every $\epsilon>0$, we have $1/|\zeta(s)| \leq c_\epsilon |t|^\epsilon$ when $s=\sigma+it$, $\sigma \geq 1$, and $|t| \geq 1$.*

Proof. From our previous observations, we clearly have that

$$|\zeta^3(\sigma)\zeta^4(\sigma+it)\zeta(\sigma+2it)| \geq 1, \quad \text{whenever } \sigma \geq 1.$$

Using the estimate for ζ in Proposition 2.7 of Chapter 6, we find that

$$|\zeta^4(\sigma+it)| \geq c|\zeta^{-3}(\sigma)|\,|t|^{-\epsilon} \geq c'(\sigma-1)^3|t|^{-\epsilon},$$

for all $\sigma \geq 1$ and $|t| \geq 1$. Thus

$$(3) \quad |\zeta(\sigma + it)| \geq c'(\sigma - 1)^{3/4}|t|^{-\epsilon/4}, \qquad \text{whenever } \sigma \geq 1 \text{ and } |t| \geq 1.$$

We now consider two separate cases, depending on whether the inequality $\sigma - 1 \geq A|t|^{-5\epsilon}$ holds, for some appropriate constant A (whose value we choose later).

If this inequality does hold, then (3) immediately provides

$$|\zeta(\sigma + it)| \geq A'|t|^{-4\epsilon},$$

and it suffices to replace 4ϵ by ϵ to conclude the proof of the desired estimate, in this case.

If, however, $\sigma - 1 < A|t|^{-5\epsilon}$, then we first select $\sigma' > \sigma$ with $\sigma' - 1 = A|t|^{-5\epsilon}$. The triangle inequality then implies

$$|\zeta(\sigma + it)| \geq |\zeta(\sigma' + it)| - |\zeta(\sigma' + it) - \zeta(\sigma + it)|,$$

and an application of the mean value theorem, together with the estimates for the derivative of ζ obtained in the previous chapter, give

$$|\zeta(\sigma' + it) - \zeta(\sigma + it)| \leq c''|\sigma' - \sigma|\,|t|^{\epsilon} \leq c''|\sigma' - 1|\,|t|^{\epsilon}.$$

These observations, together with an application of (3) where we set $\sigma = \sigma'$, show that

$$|\zeta(\sigma + it)| \geq c'(\sigma' - 1)^{3/4}|t|^{-\epsilon/4} - c''(\sigma' - 1)|t|^{\epsilon}.$$

Now choose $A = (c'/(2c''))^4$, and recall that $\sigma' - 1 = A|t|^{-5\epsilon}$. This gives precisely

$$c'(\sigma' - 1)^{3/4}|t|^{-\epsilon/4} = 2c''(\sigma' - 1)|t|^{\epsilon},$$

and therefore

$$|\zeta(\sigma + it)| \geq A''|t|^{-4\epsilon}.$$

On replacing 4ϵ by ϵ, the desired inequality is established, and the proof of the proposition is complete.

2 Reduction to the functions ψ and ψ_1

In his study of primes, Tchebychev introduced an auxiliary function whose behavior is to a large extent equivalent to the asymptotic distribution of primes, but which is easier to manipulate than $\pi(x)$. **Tchebychev's ψ-function** is defined by

$$\psi(x) = \sum_{p^m \leq x} \log p.$$

The sum is taken over those integers of the form p^m that are less than or equal to x. Here p is a prime number and m is a positive integer. There are two other formulations of ψ that we shall need. First, if we define

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^m \text{ for some prime } p \text{ and some } m \geq 1, \\ 0 & \text{otherwise,} \end{cases}$$

then it is clear that

$$\psi(x) = \sum_{1 \leq n \leq x} \Lambda(n).$$

Also, it is immediate that

$$\psi(x) = \sum_{p \leq x} \left[\frac{\log x}{\log p} \right] \log p$$

where $[u]$ denotes the greatest integer $\leq u$, and the sum is taken over the primes less than x. This formula follows from the fact that if $p^m \leq x$, then $m \leq \log x / \log p$.

The fact that $\psi(x)$ contains enough information about $\pi(x)$ to prove our theorem is given a precise meaning in the statement of the next proposition. In particular, this reduces the prime number theorem to a corresponding asymptotic statement about ψ.

Proposition 2.1 *If* $\psi(x) \sim x$ *as* $x \to \infty$, *then* $\pi(x) \sim x/\log x$ *as* $x \to \infty$.

Proof. The argument here is elementary. By definition, it suffices to prove the following two inequalities:

$$1 \leq \liminf_{x \to \infty} \pi(x) \frac{\log x}{x} \quad \text{and} \quad \limsup_{x \to \infty} \pi(x) \frac{\log x}{x} \leq 1. \tag{4}$$

To do so, first note that crude estimates give

$$\psi(x) = \sum_{p \leq x} \left[\frac{\log x}{\log p} \right] \log p \leq \sum_{p \leq x} \frac{\log x}{\log p} \log p = \pi(x) \log x,$$

and dividing through by x yields

$$\frac{\psi(x)}{x} \leq \frac{\pi(x) \log x}{x}.$$

The asymptotic condition $\psi(x) \sim x$ implies the first inequality in (4). The proof of the second inequality is a little trickier. Fix $0 < \alpha < 1$, and note that

$$\psi(x) \geq \sum_{p \leq x} \log p \geq \sum_{x^\alpha < p \leq x} \log p \geq (\pi(x) - \pi(x^\alpha)) \log x^\alpha,$$

and therefore

$$\psi(x) + \alpha\pi(x^\alpha) \log x \geq \alpha\pi(x) \log x.$$

Dividing by x, noting that $\pi(x^\alpha) \leq x^\alpha$, $\alpha < 1$, and $\psi(x) \sim x$, gives

$$1 \geq \alpha \limsup_{x\to\infty} \pi(x) \frac{\log x}{x}.$$

Since $\alpha < 1$ was arbitrary, the proof is complete.

Remark. The converse of the proposition is also true: if $\pi(x) \sim x/\log x$ then $\psi(x) \sim x$. Since we shall not need this result, we leave the proof to the interested reader.

In fact, it will be more convenient to work with a close cousin of the ψ function. Define the function ψ_1 by

$$\psi_1(x) = \int_1^x \psi(u)\, du.$$

In the previous proposition we reduced the prime number theorem to the asymptotics of $\psi(x)$ as x tends to infinity. Next, we show that this follows from the asymptotics of ψ_1.

Proposition 2.2 *If $\psi_1(x) \sim x^2/2$ as $x \to \infty$, then $\psi(x) \sim x$ as $x \to \infty$, and therefore $\pi(x) \sim x/\log x$ as $x \to \infty$.*

Proof. By Proposition 2.1, it suffices to prove that $\psi(x) \sim x$ as $x \to \infty$. This will follow quite easily from the fact that if $\alpha < 1 < \beta$, then

$$\frac{1}{(1-\alpha)x} \int_{\alpha x}^{x} \psi(u)\, du \leq \psi(x) \leq \frac{1}{(\beta-1)x} \int_{x}^{\beta x} \psi(u)\, du.$$

The proof of this double inequality is immediate and relies simply on the fact that ψ is increasing. As a consequence, we find, for example, that

$$\psi(x) \leq \frac{1}{(\beta-1)x}[\psi_1(\beta x) - \psi_1(x)],$$

and therefore

$$\frac{\psi(x)}{x} \leq \frac{1}{(\beta - 1)} \left[\frac{\psi_1(\beta x)}{(\beta x)^2} \beta^2 - \frac{\psi_1(x)}{x^2} \right].$$

In turn this implies

$$\limsup_{x \to \infty} \frac{\psi(x)}{x} \leq \frac{1}{\beta - 1} \left[\frac{1}{2} \beta^2 - \frac{1}{2} \right] = \frac{1}{2}(\beta + 1).$$

Since this result is true for all $\beta > 1$, we have proved that $\limsup_{x\to\infty} \psi(x)/x \leq 1$. A similar argument with $\alpha < 1$, then shows that $\liminf_{x\to\infty} \psi(x)/x \geq 1$, and the proof of the proposition is complete.

It is now time to relate ψ_1 (and therefore also ψ) and ζ. We proved in Lemma 1.3 that for $\mathrm{Re}(s) > 1$

$$\log \zeta(s) = \sum_{m,p} \frac{p^{-ms}}{m}.$$

Differentiating this expression gives

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{m,p} (\log p) p^{-ms} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

We record this formula for $\mathrm{Re}(s) > 1$ as

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}. \tag{5}$$

The asymptotic behavior $\psi_1(x) \sim x^2/2$ will be a consequence via (5) of the relationship between ψ_1 and ζ, which is expressed by the following noteworthy integral formula.

Proposition 2.3 *For all $c > 1$*

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) ds. \tag{6}$$

To make the proof of this formula clear, we isolate the necessary contour integrals in a lemma.

Lemma 2.4 *If $c > 0$, then*

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{a^s}{s(s+1)}\,ds = \begin{cases} 0 & \text{if } 0 < a \leq 1, \\ 1 - 1/a & \text{if } 1 \leq a. \end{cases}$$

Here, the integral is over the vertical line $\mathrm{Re}(s) = c$.

Proof. First note that since $|a^s| = a^c$, the integral converges. We suppose first that $1 \leq a$, and write $a = e^\beta$ with $\beta = \log a \geq 0$. Let

$$f(s) = \frac{a^s}{s(s+1)} = \frac{e^{s\beta}}{s(s+1)}.$$

Then $\mathrm{res}_{s=0} f = 1$ and $\mathrm{res}_{s=-1} f = -1/a$. For $T > 0$, consider the path $\Gamma(T)$ shown on Figure 1.

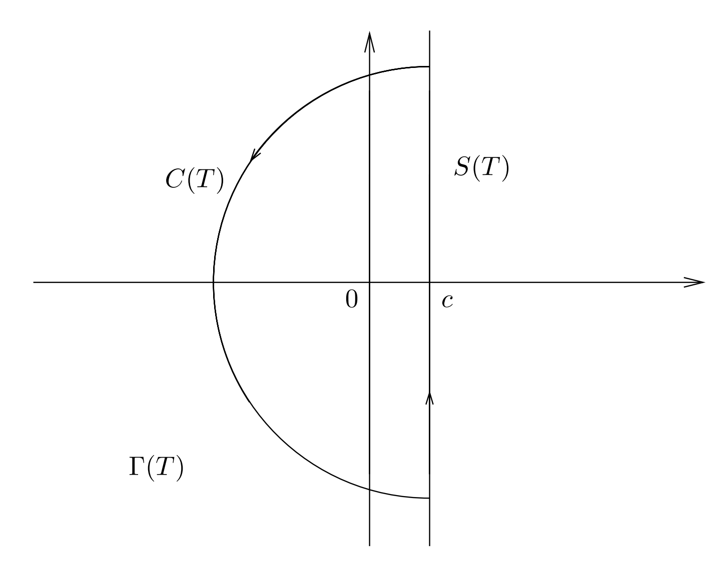


Figure 1. The contour in the proof of Lemma 2.4 when $a \geq 1$

The path $\Gamma(T)$ consists of the vertical segment $S(T)$ from $c - iT$ to $c + iT$, and of the half-circle $C(T)$ centered at c of radius T, lying to the left of the vertical segment. We equip $\Gamma(T)$ with the positive (counter-clockwise) orientation, and note that we are dealing with a toy contour. If we choose T so large that 0 and -1 are contained in the interior of $\Gamma(T)$, then by the residue formula

$$\frac{1}{2\pi i}\int_{\Gamma(T)} f(s)\,ds = 1 - 1/a.$$

Since

$$\int_{\Gamma(T)} f(s)\,ds = \int_{S(T)} f(s)\,ds + \int_{C(T)} f(s)\,ds,$$

it suffices to prove that the integral over the half-circle goes to 0 as T tends to infinity. Note that if $s = \sigma + it \in C(T)$, then for all large T we have

$$|s(s+1)| \geq (1/2)T^2,$$

and since $\sigma \leq c$ we also have the estimate $|e^{\beta s}| \leq e^{\beta c}$. Therefore

$$\left|\int_{C(T)} f(s)\,ds\right| \leq \frac{C}{T^2} 2\pi T \to 0 \qquad \text{as } T \to \infty,$$

and the case when $a \geq 1$ is proved.

If $0 < a \leq 1$, consider an analogous contour but with the half-circle lying to the right of the line $\mathrm{Re}(s) = c$. Noting that there are no poles in the interior of that contour, we can give an argument similar to the one given above to show that the integral over the half-circle also goes to 0 as T tends to infinity.

We are now ready to prove Proposition 2.3. First, observe that

$$\psi(u) = \sum_{n=1}^{\infty} \Lambda(n) f_n(u),$$

where $f_n(u) = 1$ if $n \leq u$ and $f_n(u) = 0$ otherwise. Therefore,

$$\begin{aligned}
\psi_1(x) &= \int_0^x \psi(u)\,du \\
&= \sum_{n=1}^{\infty} \int_0^x \Lambda(n) f_n(u)\,du \\
&= \sum_{n \leq x} \Lambda(n) \int_n^x du,
\end{aligned}$$

and hence

$$\psi_1(x) = \sum_{n \leq x} \Lambda(n)(x - n).$$

This fact, together with equation (5) and an application of Lemma 2.4 (with $a = x/n$), gives

$$\begin{aligned}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{x^{s+1}}{s(s+1)}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)ds &= x\sum_{n=1}^{\infty}\Lambda(n)\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{(x/n)^s}{s(s+1)}\,ds\\ &= x\sum_{n\leq x}\Lambda(n)\left(1-\frac{n}{x}\right)\\ &= \psi_1(x),\end{aligned}$$

as was to be shown.

2.1 Proof of the asymptotics for ψ_1

In this section, we will show that

$$\psi_1(x) \sim x^2/2 \quad \text{as } x \to \infty,$$

and as a consequence, we will have proved the prime number theorem.

The key ingredients in the argument are:

- the formula in Proposition 2.3 connecting ψ_1 to ζ, namely

  $$\psi_1(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{x^{s+1}}{s(s+1)}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)ds$$

  for $c > 1$.

- the non-vanishing of the zeta function on $\mathrm{Re}(s) = 1$,

  $$\zeta(1+it) \neq 0 \quad \text{for all } t \in \mathbb{R},$$

  and the estimates for ζ near that line given in Proposition 2.7 of Chapter 6 together with Proposition 1.6 of this chapter.

Let us now discuss our strategy in more detail. In the integral above for $\psi_1(x)$ we want to change the line of integration $\mathrm{Re}(s) = c$ with $c > 1$, to $\mathrm{Re}(s) = 1$. If we could achieve that, the size of the factor x^{s+1} in the integrand would then be of order x^2 (which is close to what we want) instead of x^{c+1}, $c > 1$, which is much too large. However, there would still be two issues that must be dealt with. The first is the pole of $\zeta(s)$ at $s = 1$; it turns out that when it is taken into account, its contribution is exactly the main term $x^2/2$ of the asymptotic of $\psi_1(x)$. Second, what remains must be shown to be essentially smaller than this term, and so

we must further refine the crude estimate of order x^2 when integrating on the line $\mathrm{Re}(s) = 1$. We carry out our plan as follows.

Fix $c > 1$, say $c = 2$, and assume x is also fixed for the moment with $x \geq 2$. Let $F(s)$ denote the integrand

$$F(s) = \frac{x^{s+1}}{s(s+1)} \left(-\frac{\zeta'(s)}{\zeta(s)} \right).$$

First we deform the vertical line from $c - i\infty$ to $c + i\infty$ to the path $\gamma(T)$ shown in Figure 2. (The segments of $\gamma(T)$ on the line $\mathrm{Re}(s) = 1$ consist of $T \leq t < \infty$, and $-\infty < t \leq -T$.) Here $T \geq 3$, and T will be chosen appropriately large later.

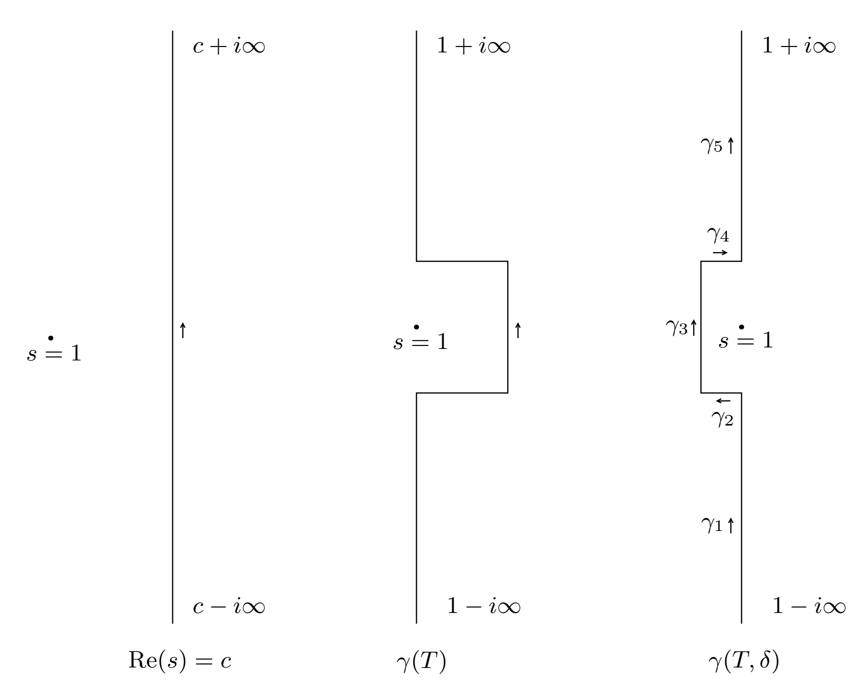


Figure 2. Three stages: the line $\mathrm{Re}(s) = c$, the contours $\gamma(T)$ and $\gamma(T, \delta)$

The usual and familiar arguments using Cauchy's theorem allow us to see that

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F(s)\, ds = \frac{1}{2\pi i} \int_{\gamma(T)} F(s)\, ds. \tag{7}$$

Indeed, we know on the basis of Proposition 2.7 in Chapter 6 and Proposition 1.6 that $|\zeta'(s)/\zeta(s)| \leq A|t|^\eta$ for any fixed $\eta > 0$, whenever $s = \sigma + it$, $\sigma \geq 1$, and $|t| \geq 1$. Thus $|F(s)| \leq A'|t|^{-2+\eta}$ in the two (infinite) rectangles bounded by the line $(c - i\infty, c + i\infty)$ and $\gamma(T)$. Since F is regular in that region, and its decrease at infinity is rapid enough, the assertion (7) is established.

Next, we pass from the contour $\gamma(T)$ to the contour $\gamma(T, \delta)$. (Again, see Figure 2.) For fixed T, we choose $\delta > 0$ small enough so that ζ has no zeros in the box

$$\{s = \sigma + it,\ 1 - \delta \leq \sigma \leq 1,\ |t| \leq T\}.$$

Such a choice can be made since ζ does not vanish on the line $\sigma = 1$.

Now $F(s)$ has a simple pole at $s = 1$. In fact, by Corollary 2.6 in Chapter 6, we know that $\zeta(s) = 1/(s-1) + H(s)$, where $H(s)$ is regular near $s = 1$. Hence $-\zeta'(s)/\zeta(s) = 1/(s-1) + h(s)$, where $h(s)$ is holomorphic near $s = 1$, and so the residue of $F(s)$ at $s = 1$ equals $x^2/2$. As a result

$$\frac{1}{2\pi i}\int_{\gamma(T)} F(s)\, ds = \frac{x^2}{2} + \frac{1}{2\pi i}\int_{\gamma(T,\delta)} \frac{x^{s+1}}{s(s+1)} F(s)\, ds.$$

We now decompose the contour $\gamma(T, \delta)$ as $\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5$ and estimate each of the integrals $\int_{\gamma_j} F(s)\, ds$, $j = 1, 2, 3, 4, 5$, with the γ_j as in Figure 2.

First we contend that there exists T so large that

$$\left|\int_{\gamma_1} F(s)\, ds\right| \leq \frac{\epsilon}{2}x^2 \quad \text{and} \quad \left|\int_{\gamma_5} F(s)\, ds\right| \leq \frac{\epsilon}{2}x^2.$$

To see this, we first note that for $s \in \gamma_1$ one has

$$|x^{1+s}| = x^{1+\sigma} = x^2.$$

Then, by Proposition 1.6 we have, for example, that $|\zeta'(s)/\zeta(s)| \leq A|t|^{1/2}$, so

$$\left|\int_{\gamma_1} F(s)\, ds\right| \leq Cx^2 \int_T^\infty \frac{|t|^{1/2}}{t^2}\, dt.$$

Since the integral converges, we can make the right-hand side $\leq \epsilon x^2/2$ upon taking T sufficiently large. The argument for the integral over γ_5 is the same.

Having now fixed T, we choose δ appropriately small. On γ_3, note that

$$|x^{1+s}| = x^{1+1-\delta} = x^{2-\delta}\,,$$

from which we conclude that there exists a constant C_T (dependent on T) such that

$$\left|\int_{\gamma_3} F(s)\,ds\right| \leq C_T x^{2-\delta}\,.$$

Finally, on the small horizontal segment γ_2 (and similarly on γ_4), we can estimate the integral as follows:

$$\left|\int_{\gamma_2} F(s)\,ds\right| \leq C'_T \int_{1-\delta}^{1} x^{1+\sigma}\,d\sigma \leq C'_T \frac{x^2}{\log x}\,.$$

We conclude that there exist constants C_T and C'_T (possibly different from the ones above) such that

$$\left|\psi_1(x) - \frac{x^2}{2}\right| \leq \epsilon x^2 + C_T x^{2-\delta} + C'_T \frac{x^2}{\log x}.$$

Dividing through by $x^2/2$, we see that

$$\left|\frac{2\psi_1(x)}{x^2} - 1\right| \leq 2\epsilon + 2C_T x^{-\delta} + 2C'_T \frac{1}{\log x},$$

and therefore, for all large x we have

$$\left|\frac{2\psi_1(x)}{x^2} - 1\right| \leq 4\epsilon.$$

This concludes the proof that

$$\psi_1(x) \sim x^2/2 \quad \text{ as } x \to \infty\,,$$

and thus, we have also completed the proof of the prime number theorem.

Note on interchanging double sums

We prove the following facts about the interchange of infinite sums: if $\{a_{k\ell}\}_{1\leq k,\ell<\infty}$ is a sequence of complex numbers indexed by $\mathbb{N}\times\mathbb{N}$, such that

$$\sum_{k=1}^{\infty}\left(\sum_{\ell=1}^{\infty} |a_{k\ell}|\right) < \infty, \tag{8}$$

then:

(i) The double sum $A = \sum_{k=1}^{\infty}\left(\sum_{\ell=1}^{\infty} a_{k\ell}\right)$ summed in this order converges, and we may in fact also interchange the order of summation, so that

$$A = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} a_{k\ell} = \sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty} a_{k\ell}.$$

(ii) Given $\epsilon > 0$, there is a positive integer N so that for all $K, L > N$ we have $\left|A - \sum_{k=1}^{K}\sum_{\ell=1}^{L} a_{k\ell}\right| < \epsilon$.

(iii) If $m \mapsto (k(m), \ell(m))$ is a bijection from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$, and if we write $c_m = a_{k(m)\ell(m)}$, then $A = \sum_{k=1}^{\infty} c_k$.

Statement (iii) says that any rearrangement of the sequence $\{a_{k\ell}\}$ can be summed without changing the limit. This is analogous to the case of absolutely convergent series, which can be summed in any desired order.

The condition (8) says that each sum $\sum_{\ell} a_{k\ell}$ converges absolutely, and moreover this convergence is "uniform" in k. An analogous situation arises for sequences of functions, where an important question is whether or not the interchange of limits

$$\lim_{x\to x_0}\lim_{n\to\infty} f_n(x) \stackrel{?}{=} \lim_{n\to\infty}\lim_{x\to x_0} f_n(x)$$

holds. It is a well-known fact that if the f_n's are continuous, and their convergence is uniform, then the above identity is true since the limit function is itself continuous. To take advantage of this fact, define $b_k = \sum_{\ell=1}^{\infty} |a_{k\ell}|$ and let $S = \{x_0, x_1, \ldots\}$ be a countable set of points with $\lim_{n\to\infty} x_n = x_0$. Also, define functions on S as follows:

$$\begin{aligned} f_k(x_0) &= \textstyle\sum_{\ell=1}^{\infty} a_{k\ell} && \text{for } k = 1, 2, \ldots \\ f_k(x_n) &= \textstyle\sum_{\ell=1}^{n} a_{k\ell} && \text{for } k = 1, 2, \ldots \text{ and } n = 1, 2, \ldots \\ g(x) &= \textstyle\sum_{k=1}^{\infty} f_k(x) && \text{for } x \in S. \end{aligned}$$

By assumption (8), each f_k is continuous at x_0. Moreover $|f_k(x)| \le b_k$ and $\sum b_k < \infty$, so the series defining the function g is uniformly convergent on S, and therefore g is also continuous at x_0. As a consequence we find (i), since

$$\begin{aligned} \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} a_{k\ell} = g(x_0) = \lim_{n\to\infty} g(x_n) &= \lim_{n\to\infty}\sum_{k=1}^{\infty}\sum_{\ell=1}^{n} a_{k\ell} \\ = \lim_{n\to\infty}\sum_{\ell=1}^{n}\sum_{k=1}^{\infty} a_{k\ell} &= \sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty} a_{k\ell}. \end{aligned}$$

For the second statement, first observe that

$$\left|A - \sum_{k=1}^{K}\sum_{\ell=1}^{L} a_{k\ell}\right| \le \sum_{k\le K}\sum_{\ell>L} |a_{k\ell}| + \sum_{k>K}\sum_{\ell=1}^{\infty} |a_{k\ell}|.$$

To estimate the second term, we use the fact that $\sum b_k$ converges, which implies $\sum_{k>K}\sum_{\ell=1}^{\infty}|a_{k\ell}|<\epsilon/2$ whenever $K>K_0$, for some K_0. For the first term above, note that $\sum_{k\le K}\sum_{\ell>L}|a_{k\ell}|\le\sum_{k=1}^{\infty}\sum_{\ell>L}|a_{k\ell}|$. But the argument above guarantees that we can interchange these last two sums; also $\sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty}|a_{k\ell}|<\infty$, so that for all $L>L_0$ we have $\sum_{\ell>L}\sum_{k=1}^{\infty}|a_{k\ell}|<\epsilon/2$. Taking $N>\max(L_0,K_0)$ completes the proof of (ii).

The proof of (iii) is a direct consequence of (ii). Indeed, given any rectangle

$$R(K,L)=\{(k,\ell)\in\mathbb{N}\times\mathbb{N}:1\le k\le K\text{ and }1\le\ell\le L\},$$

there exists M such that the image of $[1,M]$ under the map $m\mapsto(k(m),\ell(m))$ contains $R(K,L)$.

When U denotes any open set in $\mathbb{R}^2$ that contains the origin, we define for $R>0$ its dilate $U(R)=\{y\in\mathbb{R}^2: y=Rx\ \text{ for some } x\in U\}$, and we can apply (ii) to see that

$$A=\lim_{R\to\infty}\sum_{(k,\ell)\in U(R)}a_{k\ell}.$$

In other words, under condition (8) the double sum $\sum_{k\ell}a_{k\ell}$ can be evaluated by summing over discs, squares, rectangles, ellipses, etc.

Finally, we leave the reader with the instructive task of finding a sequence of complex numbers $\{a_{k\ell}\}$ such that

$$\sum_k\sum_\ell a_{k\ell}\neq\sum_\ell\sum_k a_{k\ell}.$$

[Hint: Consider $\{a_{k\ell}\}$ as the entries of an infinite matrix with 0 above the diagonal, -1 on the diagonal, and $a_{k\ell}=2^{\ell-k}$ if $k>\ell$.]

3 Exercises

1. Suppose that $\{a_n\}_{n=1}^{\infty}$ is a sequence of real numbers such that the partial sums

$$A_n=a_1+\cdots+a_n$$

are bounded. Prove that the Dirichlet series

$$\sum_{n=1}^{\infty}\frac{a_n}{n^s}$$

converges for $\mathrm{Re}(s)>0$ and defines a holomorphic function in this half-plane.

[Hint: Use summation by parts to compare the original (non-absolutely convergent) series to the (absolutely convergent) series $\sum A_n(n^{-s}-(n+1)^{-s})$. An estimate for the term in parentheses is provided by the mean value theorem. To prove that the series is analytic, show that the partial sums converge uniformly on every compact subset of the half-plane $\mathrm{Re}(s)>0$.]

2. The following links the multiplication of Dirichlet series with the divisibility properties of their coefficients.

(a) Show that if $\{a_m\}$ and $\{b_k\}$ are two bounded sequences of complex numbers, then

$$\left(\sum_{m=1}^{\infty} \frac{a_m}{m^s}\right)\left(\sum_{k=1}^{\infty} \frac{b_k}{k^s}\right) = \sum_{n=1}^{\infty} \frac{c_n}{n^s} \qquad \text{where } c_n = \sum_{mk=n} a_m b_k.$$

The above series converge absolutely when $\text{Re}(s) > 1$.

(b) Prove as a consequence that one has

$$(\zeta(s))^2 = \sum_{n=1}^{\infty} \frac{d(n)}{n^s} \qquad \text{and} \qquad \zeta(s)\zeta(s-a) = \sum_{n=1}^{\infty} \frac{\sigma_a(n)}{n^s}$$

for $\text{Re}(s) > 1$ and $\text{Re}(s-a) > 1$, respectively. Here $d(n)$ equals the number of divisors of n, and $\sigma_a(n)$ is the sum of the ath powers of divisors of n. In particular, one has $\sigma_0(n) = d(n)$.

3. In line with the previous exercise, we consider the Dirichlet series for $1/\zeta$.

(a) Prove that for $\text{Re}(s) > 1$

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s},$$

where $\mu(n)$ is the **Möbius function** defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1\,, \\ (-1)^k & \text{if } n = p_1 \cdots p_k, \text{ and the } p_j \text{ are distinct primes}\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Note that $\mu(nm) = \mu(n)\mu(m)$ whenever n and m are relatively prime. [Hint: Use the Euler product formula for $\zeta(s)$.]

(b) Show that

$$\sum_{k|n} \mu(k) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

4. Suppose $\{a_n\}_{n=1}^{\infty}$ is a sequence of complex numbers such that $a_n = a_m$ if $n \equiv m \mod q$ for some positive integer q. Define the **Dirichlet L-series** associated to $\{a_n\}$ by

$$L(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} \qquad \text{for } \text{Re}(s) > 1.$$

Also, with $a_0 = a_q$, let

$$Q(x) = \sum_{m=0}^{q-1} a_{q-m} e^{mx}.$$

Show, as in Exercises 15 and 16 of the previous chapter, that

$$L(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{Q(x)x^{s-1}}{e^{qx}-1}\, dx, \qquad \text{for } \mathrm{Re}(s) > 1.$$

Prove as a result that $L(s)$ is continuable into the complex plane, with the only possible singularity a pole at $s = 1$. In fact, $L(s)$ is regular at $s = 1$ if and only if $\sum_{m=0}^{q-1} a_m = 0$. Note the connection with the Dirichlet $L(s, \chi)$ series, taken up in Book I, Chapter 8, and that as a consequence, $L(s, \chi)$ is regular at $s = 1$ if and only if χ is a non-trivial character.

5. Consider the following function

$$\tilde{\zeta}(s) = 1 - \frac{1}{2^s} + \frac{1}{3^s} - \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s}.$$

(a) Prove that the series defining $\tilde{\zeta}(s)$ converges for $\mathrm{Re}(s) > 0$ and defines a holomorphic function in that half-plane.

(b) Show that for $s > 1$ one has $\tilde{\zeta}(s) = (1 - 2^{1-s})\zeta(s)$.

(c) Conclude, since $\tilde{\zeta}$ is given as an alternating series, that ζ has no zeros on the segment $0 < \sigma < 1$. Extend this last assertion to $\sigma = 0$ by using the functional equation.

6. Show that for every $c > 0$

$$\lim_{N\to\infty} \frac{1}{2\pi i} \int_{c-iN}^{c+iN} a^s \frac{ds}{s} = \begin{cases} 1 & \text{if } a > 1, \\ 1/2 & \text{if } a = 1, \\ 0 & \text{if } 0 \le a < 1. \end{cases}$$

The integral is taken over the vertical segment from $c - iN$ to $c + iN$.

7. Show that the function

$$\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$$

is real when s is real, or when $\mathrm{Re}(s) = 1/2$.

8. The function ζ has infinitely many zeros in the critical strip. This can be seen as follows.

(a) Let

$$F(s) = \xi(1/2 + s), \qquad \text{where} \qquad \xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s).$$

Show that $F(s)$ is an even function of s, and as a result, there exists G so that $G(s^2) = F(s)$.

(b) Show that the function $(s-1)\zeta(s)$ is an entire function of growth order 1, that is

$$|(s-1)\zeta(s)| \leq A_\epsilon e^{a_\epsilon |s|^{1+\epsilon}}.$$

As a consequence $G(s)$ is of growth order 1/2.

(c) Deduce from the above that ζ has infinitely many zeros in the critical strip.

[Hint: To prove (a) and (b) use the functional equation for $\zeta(s)$. For (c), use a result of Hadamard, which states that an entire function with fractional order has infinitely many zeros (Exercise 14 in Chapter 5).]

9. Refine the estimates in Proposition 2.7 in Chapter 6 and Proposition 1.6 to show that

(a) $|\zeta(1 + it)| \leq A \log |t|$,

(b) $|\zeta'(1 + it)| \leq A(\log |t|)^2$,

(c) $1/|\zeta(1 + it)| \leq A(\log |t|)^a$,

when $|t| \geq 2$ (with $a = 7$).

10. In the theory of primes, a better approximation to $\pi(x)$ (instead of $x/\log x$) turns out to be $\mathrm{Li}(x)$ defined by

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}.$$

(a) Prove that

$$\mathrm{Li}(x) = \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right) \qquad \text{as } x \to \infty,$$

and that as a consequence

$$\pi(x) \sim \mathrm{Li}(x) \qquad \text{as } x \to \infty.$$

[Hint: Integrate by parts in the definition of $\mathrm{Li}(x)$ and observe that it suffices to prove

$$\int_2^x \frac{dt}{(\log t)^2} = O\left(\frac{x}{(\log x)^2}\right).$$

To see this, split the integral from 2 to $\sqrt{x}$ and from $\sqrt{x}$ to x.]

(b) Refine the previous analysis by showing that for every integer $N > 0$ one has the following asymptotic expansion

$$\mathrm{Li}(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + 2\frac{x}{(\log x)^3} \cdots + (N-1)!\frac{x}{(\log x)^N} + O\left(\frac{x}{(\log x)^{N+1}}\right)$$

as $x \to \infty$.

11. Let

$$\varphi(x) = \sum_{p \leq x} \log p$$

where the sum is taken over all primes $\leq x$. Prove that the following are equivalent as $x \to \infty$:

(i) $\varphi(x) \sim x$,

(ii) $\pi(x) \sim x/\log x$,

(iii) $\psi(x) \sim x$,

(iv) $\psi_1(x) \sim x^2/2$.

12. If p_n denotes the nth prime, the prime number theorem implies that $p_n \sim n \log n$ as $n \to \infty$.

(a) Show that $\pi(x) \sim x/\log x$ implies that

$$\log \pi(x) + \log \log x \sim \log x.$$

(b) As a consequence, prove that $\log \pi(x) \sim \log x$, and take $x = p_n$ to conclude the proof.

4 Problems

1. Let $F(s) = \sum_{n=1}^{\infty} a_n/n^s$, where $|a_n| \leq M$ for all n.

(a) Then

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |F(\sigma + it)|^2 \, dt = \sum_{n=1}^{\infty} \frac{|a_n|^2}{n^{2\sigma}} \qquad \text{if } \sigma > 1.$$

How is this reminiscent of the Parseval-Plancherel theorem? See e.g. Chapter 3 in Book I.

(b) Show as a consequence the uniqueness of Dirichlet series: If $F(s)=\sum_{n=1}^{\infty} a_n n^{-s}$, where the coefficients are assumed to satisfy $|a_n| \leq cn^k$ for some k, and $F(s) \equiv 0$, then $a_n = 0$ for all n.

Hint: For part (a) use the fact that

$$\frac{1}{2T} \int_{-T}^{T} (nm)^{-\sigma} n^{-it} m^{it} \, dt \to \begin{cases} n^{-2\sigma} & \text{if } n = m, \\ 0 & \text{if } n \neq m. \end{cases}$$

2.* One of the "explicit formulas" in the theory of primes is as follows: if ψ_1 is the integrated Tchebychev function considered in Section 2, then

$$\psi_1(x) = \frac{x^2}{2} - \sum_{\rho} \frac{x^{\rho}}{\rho(\rho+1)} - E(x)$$

where the sum is taken over all zeros ρ of the zeta function in the critical strip. The error term is given by $E(x) = c_1 x + c_0 + \sum_{k=1}^{\infty} x^{1-2k}/(2k(2k-1))$, where $c_1 = \zeta'(0)/\zeta(0)$ and $c_0 = \zeta'(-1)/\zeta(-1)$. Note that $\sum_{\rho} 1/|\rho|^{1+\epsilon} < \infty$ for every $\epsilon > 0$, because $(1-s)\zeta(s)$ has order of growth 1. (See Exercise 8.) Also, obviously $E(x) = O(x)$ as $x \to \infty$.

3.* Using the previous problem one can show that

$$\pi(x) - \mathrm{Li}(x) = O(x^{\alpha+\epsilon}) \qquad \text{as } x \to \infty$$

for every $\epsilon > 0$, where α is fixed and $1/2 \leq \alpha < 1$ if and only if $\zeta(s)$ has no zeros in the strip $\alpha < \mathrm{Re}(s) < 1$. The case $\alpha = 1/2$ corresponds to the Riemann hypothesis.

4.* One can combine ideas from the prime number theorem with the proof of Dirichlet's theorem about primes in arithmetic progression (given in Book I) to prove the following. Let q and ℓ be relatively prime integers. We consider the primes belonging to the arithmetic progression $\{qk+\ell\}_{k=1}^{\infty}$, and let $\pi_{q,\ell}(x)$ denote the number of such primes $\leq x$. Then one has

$$\pi_{q,\ell}(x) \sim \frac{x}{\varphi(q) \log x} \qquad \text{as } x \to \infty,$$

where $\varphi(q)$ denotes the number of positive integers less than q and relatively prime to q.

8 Conformal Mappings

> The results I found for polygons can be extended under very general assumptions. I have undertaken this research because it is a step towards a deeper understanding of the mapping problem, for which not much has happened since Riemann's inaugural dissertation; this, even though the theory of mappings, with its close connection with the fundamental theorems of Riemann's function theory, deserves in the highest degree to be developed further.
>
> *E. B. Christoffel,* 1870

The problems and ideas we present in this chapter are more geometric in nature than the ones we have seen so far. In fact, here we will be primarily interested in mapping properties of holomorphic functions. In particular, most of our results will be "global," as opposed to the more "local" analytical results proved in the first three chapters. The motivation behind much of our presentation lies in the following simple question:

> Given two open sets U and V in $\mathbb{C}$, does there exist a holomorphic bijection between them?

By a holomorphic bijection we simply mean a function that is both holomorphic and bijective. (It will turn out that the inverse map is then automatically holomorphic.) A solution to this problem would permit a transfer of questions about analytic functions from one open set with little geometric structure to another with possibly more useful properties. The prime example consists in taking $V = \mathbb{D}$ the unit disc, where many ideas have been developed to study analytic functions.[1] In fact, since the disc seems to be the most fruitful choice for V we are led to a variant of the above question:

> Given an open subset Ω of $\mathbb{C}$, what conditions on Ω guarantee that there exists a holomorphic bijection from Ω to $\mathbb{D}$?

---

[1]For the corresponding problem when $V = \mathbb{C}$, the solution is trivial: only $U = \mathbb{C}$ is possible. See Exercise 14 in Chapter 3.

In some instances when a bijection exists it can be given by explicit formulas, and we turn to this aspect of the theory first. For example, the upper half-plane can be mapped by a holomorphic bijection to the disc, and this is given by a fractional linear transformation. From there, one can construct many other examples, by composing simple maps already encountered earlier, such as rational functions, trigonometric functions, logarithms, etc. As an application, we discuss the consequence of these constructions to the solution of the Dirichlet problem for the Laplacian in some particular domains.

Next, we pass from the specific examples to prove the first general result of the chapter, namely the Schwarz lemma, with an immediate application to the determination of all holomorphic bijections ("automorphisms" of the disc to itself). These are again given by fractional linear transformations.

Then comes the heart of the matter: the Riemann mapping theorem, which states that Ω can be mapped to the unit disc whenever it is simply connected and not all of $\mathbb{C}$. This is a remarkable theorem, since little is assumed about Ω, not even regularity of its boundary $\partial\Omega$. (After all, the boundary of the disc is smooth.) In particular, the interiors of triangles, squares, and in fact any polygon can be mapped via a bijective holomorphic function to the disc. A precise description of the mapping in the case of polygons, called the Schwarz-Christoffel formula, will be taken up in the last section of the chapter. It is interesting to note that the mapping functions for rectangles are given by "elliptic integrals," and these lead to doubly-periodic functions. The latter are the subject of the next chapter.

1 Conformal equivalence and examples

We fix some terminology that we shall use in the rest of this chapter. A bijective holomorphic function $f : U \to V$ is called a **conformal map** or **biholomorphism**. Given such a mapping f, we say that U and V are **conformally equivalent** or simply **biholomorphic**. An important fact is that the inverse of f is then automatically holomorphic.

Proposition 1.1 *If $f : U \to V$ is holomorphic and injective, then $f'(z) \neq 0$ for all $z \in U$. In particular, the inverse of f defined on its range is holomorphic, and thus the inverse of a conformal map is also holomorphic.*

Proof. We argue by contradiction, and suppose that $f'(z_0) = 0$ for some $z_0 \in U$. Then

$$f(z) - f(z_0) = a(z - z_0)^k + G(z) \quad \text{for all } z \text{ near } z_0,$$

with $a \neq 0$, $k \geq 2$ and G vanishing to order $k + 1$ at z_0. For sufficiently small w, we write

$$f(z) - f(z_0) - w = F(z) + G(z), \quad \text{where } F(z) = a(z - z_0)^k - w.$$

Since $|G(z)| < |F(z)|$ on a small circle centered at z_0, and F has at least two zeros inside that circle, Rouché's theorem implies that $f(z) - f(z_0) - w$ has at least two zeros there. Since $f'(z) \neq 0$ for all $z \neq z_0$ but sufficiently close to z_0 it follows that the roots of $f(z) - f(z_0) - w$ are distinct, hence f is not injective, a contradiction.

Now let $g = f^{-1}$ denote the inverse of f on its range, which we can assume is V. Suppose $w_0 \in V$ and w is close to w_0. Write $w = f(z)$ and $w_0 = f(z_0)$. If $w \neq w_0$, we have

$$\frac{g(w) - g(w_0)}{w - w_0} = \frac{1}{\frac{w - w_0}{g(w) - g(w_0)}} = \frac{1}{\frac{f(z) - f(z_0)}{z - z_0}}.$$

Since $f'(z_0) \neq 0$, we may let $z \to z_0$ and conclude that g is holomorphic at w_0 with $g'(w_0) = 1/f'(g(w_0))$.

From this proposition we conclude that two open sets U and V are conformally equivalent if and only if there exist holomorphic functions $f : U \to V$ and $g : V \to U$ such that $g(f(z)) = z$ and $f(g(w)) = w$ for all $z \in U$ and $w \in V$.

We point out that the terminology adopted here is not universal. Some authors call a holomorphic map $f : U \to V$ conformal if $f'(z) \neq 0$ for all $z \in U$. This definition is clearly less restrictive than ours; for example, $f(z) = z^2$ on the punctured disc $\mathbb{C} - \{0\}$ satisfies $f'(z) \neq 0$, but is not injective. However, the condition $f'(z) \neq 0$ is tantamount to f being a local bijection (Exercise 1). There is a geometric consequence of the condition $f'(z) \neq 0$ and it is at the root of this discrepency of terminology in the definitions. A holomorphic map that satisfies this condition preserves angles. Loosely speaking, if two curves γ and η intersect at z_0, and α is the oriented angle between the tangent vectors to these curves, then the image curves $f \circ \gamma$ and $f \circ \eta$ intersect at $f(z_0)$, and their tangent vectors form the same angle α. Problem 2 develops this idea.

We begin our study of conformal mappings by looking at a number of specific examples. The first gives the conformal equivalence between

the unit disc and the upper half-plane, which plays an important role in many problems.

1.1 The disc and upper half-plane

The upper half-plane, which we denote by $\mathbb{H}$, consists of those complex numbers with positive imaginary part; that is,

$$\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}.$$

A remarkable fact, which at first seems surprising, is that the unbounded set $\mathbb{H}$ is conformally equivalent to the unit disc. Moreover, an explicit formula giving this equivalence exists. Indeed, let

$$F(z) = \frac{i-z}{i+z} \quad \text{and} \quad G(w) = i\frac{1-w}{1+w}.$$

Theorem 1.2 *The map $F : \mathbb{H} \to \mathbb{D}$ is a conformal map with inverse $G : \mathbb{D} \to \mathbb{H}$.*

Proof. First we observe that both maps are holomorphic in their respective domains. Then we note that any point in the upper half-plane is closer to i than to $-i$, so $|F(z)| < 1$ and F maps $\mathbb{H}$ into $\mathbb{D}$. To prove that G maps into the upper half-plane, we must compute $\mathrm{Im}(G(w))$ for $w \in \mathbb{D}$. To this end we let $w = u + iv$, and note that

$$\begin{aligned}
\mathrm{Im}(G(w)) &= \mathrm{Re}\left(\frac{1-u-iv}{1+u+iv}\right) \\
&= \mathrm{Re}\left(\frac{(1-u-iv)(1+u-iv)}{(1+u)^2+v^2}\right) \\
&= \frac{1-u^2-v^2}{(1+u)^2+v^2} > 0
\end{aligned}$$

since $|w| < 1$. Therefore G maps the unit disc to the upper half-plane. Finally,

$$F(G(w)) = \frac{i - i\frac{1-w}{1+w}}{i + i\frac{1-w}{1+w}} = \frac{1+w-1+w}{1+w+1-w} = w,$$

and similarly $G(F(z)) = z$. This proves the theorem.

An interesting aspect of these functions is their behavior on the boundaries of our open sets.[2] Observe that F is holomorphic everywhere on $\mathbb{C}$

[2]The boundary behavior of conformal maps is a recurrent theme that plays an important role in this chapter.

except at $z = -i$, and in particular it is continuous everywhere on the boundary of $\mathbb{H}$, namely the real line. If we take $z = x$ real, then the distance from x to i is the same as the distance from x to $-i$, therefore $|F(x)| = 1$. Thus F maps $\mathbb{R}$ onto the boundary of $\mathbb{D}$. We get more information by writing

$$F(x) = \frac{i-x}{i+x} = \frac{1-x^2}{1+x^2} + i\frac{2x}{1+x^2},$$

and parametrizing the real line by $x = \tan t$ with $t \in (-\pi/2, \pi/2)$. Since

$$\sin 2a = \frac{2\tan a}{1+\tan^2 a} \quad \text{and} \quad \cos 2a = \frac{1-\tan^2 a}{1+\tan^2 a},$$

we have $F(x) = \cos\, 2t + i \sin\, 2t = e^{i2t}$. Hence the image of the real line is the arc consisting of the circle omitting the point -1. Moreover, as x travels from $-\infty$ to ∞, $F(x)$ travels along that arc starting from -1 and first going through that part of the circle that lies in the lower half-plane.

The point -1 on the circle corresponds to the "point at infinity" of the upper half-plane.

Remark. Mappings of the form

$$z \mapsto \frac{az+b}{cz+d},$$

where a, b, c, and d are complex numbers, and where the denominator is assumed not to be a multiple of the numerator, are usually referred to as **fractional linear transformations**. Other instances occur as the automorphisms of the disc and of the upper half-plane in Theorems 2.1 and 2.4.

1.2 Further examples

We gather here several illustrations of conformal mappings. In certain cases we discuss the behavior of the map on the boundary of the relevant domain. Some of the mappings are pictured in Figure 1.

Example 1. Translations and dilations provide the first simple examples. Indeed, if $h \in \mathbb{C}$, the translation $z \mapsto z + h$ is a conformal map from $\mathbb{C}$ to itself whose inverse is $w \mapsto w - h$. If h is real, then this translation is also a conformal map from the upper half-plane to itself.

For any non-zero complex number c, the map $f : z \mapsto cz$ is a conformal map from the complex plane to itself, whose inverse is simply $g : w \mapsto c^{-1}w$. If c has modulus 1, so that $c = e^{i\varphi}$ for some real φ, then f is

a **rotation** by φ. If $c > 0$ then f corresponds to a dilation. Finally, if $c < 0$ the map f consists of a dilation by $|c|$ followed by a rotation of π.

EXAMPLE 2. If n is a positive integer, then the map $z \mapsto z^n$ is conformal from the sector $S = \{z \in \mathbb{C} : 0 < \arg(z) < \pi/n\}$ to the upper half-plane. The inverse of this map is simply $w \mapsto w^{1/n}$, defined in terms of the principal branch of the logarithm.

More generally, if $0 < \alpha < 2$ the map $f(z) = z^\alpha$ takes the upper half-plane to the sector $S = \{w \in \mathbb{C} :\ 0 < \arg(w) < \alpha\pi\}$. Indeed, if we choose the branch of the logarithm obtained by deleting the positive real axis, and $z = re^{i\theta}$ with $r > 0$ and $0 < \theta < \pi$, then

$$f(z) = z^\alpha = |z|^\alpha e^{i\alpha\theta}.$$

Therefore f maps $\mathbb{H}$ into S. Moreover, a simple verification shows that the inverse of f is given by $g(w) = w^{1/\alpha}$, where the branch of the logarithm is chosen so that $0 < \arg w < \alpha\pi$.

By composing the map just discussed with the translations and rotations in the previous example, we may map the upper half-plane conformally to any (infinite) sector in $\mathbb{C}$.

Let us note the boundary behavior of f. If x travels from $-\infty$ to 0 on the real line, then $f(x)$ travels from $\infty e^{i\alpha\pi}$ to 0 on the half-line determined by $\arg z = \alpha\pi$. As x goes from 0 to ∞ on the real line, the image $f(x)$ goes from 0 to ∞ on the real line as well.

EXAMPLE 3. The map $f(z) = (1+z)/(1-z)$ takes the upper half-disc $\{z = x + iy :\ |z| < 1 \text{ and } y > 0\}$ conformally to the first quadrant $\{w = u + iv :\ u > 0 \text{ and } v > 0\}$. Indeed, if $z = x + iy$ we have

$$f(z) = \frac{1-(x^2+y^2)}{(1-x)^2+y^2} + i\frac{2y}{(1-x)^2+y^2},$$

so f maps the half-disc in the upper half-plane into the first quadrant. The inverse map, given by $g(w) = (w-1)/(w+1)$, is clearly holomorphic in the first quadrant. Moreover, $|w+1| > |w-1|$ for all w in the first quadrant because the distance from w to -1 is greater than the distance from w to 1; thus g maps into the unit disc. Finally, an easy calculation shows that the imaginary part of $g(w)$ is positive whenever w is in the first quadrant. So g transforms the first quadrant into the desired half-disc and we conclude that f is conformal because g is the inverse of f.

To examine the action of f on the boundary, note that if $z = e^{i\theta}$ be-

longs to the upper half-circle, then

$$f(z) = \frac{1+e^{i\theta}}{1-e^{i\theta}} = \frac{e^{-i\theta/2}+e^{i\theta/2}}{e^{-i\theta/2}-e^{i\theta/2}} = \frac{i}{\tan(\theta/2)}.$$

As θ travels from 0 to π we see that $f(e^{i\theta})$ travels along the imaginary axis from infinity to 0. Moreover, if $z = x$ is real, then

$$f(z) = \frac{1+x}{1-x}$$

is also real; and one sees from this, that f is actually a bijection from $(-1, 1)$ to the positive real axis, with $f(x)$ increasing from 0 to infinity as x travels from -1 to 1. Note also that $f(0) = 1$.

EXAMPLE 4. The map $z \mapsto \log z$, defined as the branch of the logarithm obtained by deleting the negative imaginary axis, takes the upper half-plane to the strip $\{w = u + iv : u \in \mathbb{R},\ 0 < v < \pi\}$. This is immediate from the fact that if $z = re^{i\theta}$ with $-\pi/2 < \theta < 3\pi/2$, then by definition,

$$\log z = \log r + i\theta.$$

The inverse map is then $w \mapsto e^w$.

As x travels from $-\infty$ to 0, the point $f(x)$ travels from $\infty + i\pi$ to $-\infty + i\pi$ on the line $\{x + i\pi : -\infty < x < \infty\}$. When x travels from 0 to ∞ on the real line, its image $f(x)$ then goes from $-\infty$ to ∞ along the reals.

EXAMPLE 5. With the previous example in mind, we see that $z \mapsto \log z$ also defines a conformal map from the half-disc $\{z = x + iy : |z| < 1,\ y > 0\}$ to the half-strip $\{w = u + iv : u < 0,\ 0 < v < \pi\}$. As x travels from 0 to 1 on the real line, then $\log x$ goes from $-\infty$ to 0. When x goes from 1 to -1 on the half-circle in the upper half-plane, then the point $\log x$ travels from 0 to πi on the vertical segment of the strip. Finally, as x goes from -1 to 0, the point $\log x$ goes from πi to $-\infty + i\pi$ on the top half-line of the strip.

EXAMPLE 6. The map $f(z) = e^{iz}$ takes the half-strip $\{z = x + iy : -\pi/2 < x < \pi/2,\ \ y > 0\}$ conformally to the half-disc $\{w = u + iv : |w| < 1,\ u > 0\}$. This is immediate from the fact that if $z = x + iy$, then

$$e^{iz} = e^{-y}e^{ix}.$$

If x goes from $\pi/2 + i\infty$ to $\pi/2$, then $f(x)$ goes from 0 to i, and as x goes from $\pi/2$ to $-\pi/2$, then $f(x)$ travels from i to $-i$ on the half-circle. Finally, as x goes from $-\pi/2$ to $-\pi/2 + i\infty$, we see that $f(x)$ travels from $-i$ back to 0.

The mapping f is closely related to the inverse of the map in Example 5.

Example 7. The function $f(z) = -\frac{1}{2}(z + 1/z)$ is a conformal map from the half-disc $\{z = x + iy : |z| < 1,\ y > 0\}$ to the upper half-plane (Exercise 5).

The boundary behavior of f is as follows. If x travels from 0 to 1, then $f(x)$ goes from ∞ to 1 on the real axis. If $z = e^{i\theta}$, then $f(z) = \cos\theta$ and as x travels from 1 to -1 along the unit half-circle in the upper half-plane, the $f(x)$ goes from 1 to -1 on the real segment. Finally, when x goes from -1 to 0, $f(x)$ goes from -1 to $-\infty$ along the real axis.

Example 8. The map $f(z) = \sin z$ takes the upper half-plane conformally onto the half-strip $\{w = x + iy : -\pi/2 < x < \pi/2\ y > 0\}$. To see this, note that if $\zeta = e^{iz}$, then

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} = \frac{-1}{2}\left(i\zeta + \frac{1}{i\zeta}\right),$$

and therefore f is obtained first by applying the map in Example 6, then multiplying by i (that is, rotating by $\pi/2$), and finally applying the map in Example 7.

As x travels from $-\pi/2 + i\infty$ to $-\pi/2$, the point $f(x)$ goes from $-\infty$ to -1. When x is real, between $-\pi/2$ and $\pi/2$, then $f(x)$ is also real between -1 and 1. Finally, if x goes from $\pi/2$ to $\pi/2 + i\infty$, then $f(x)$ travels from 1 to ∞ on the real axis.

1.3 The Dirichlet problem in a strip

The Dirichlet problem in the open set Ω consists of solving

$$\begin{cases} \triangle u &= 0 \quad \text{in } \Omega, \\ u &= f \quad \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $\triangle$ denotes the Laplacian $\partial^2/\partial x^2 + \partial^2/\partial y^2$, and f is a given function on the boundary of Ω. In other words, we wish to find a harmonic function in Ω with prescribed boundary values f. This problem was already considered in Book I in the cases where Ω is the unit disc or the

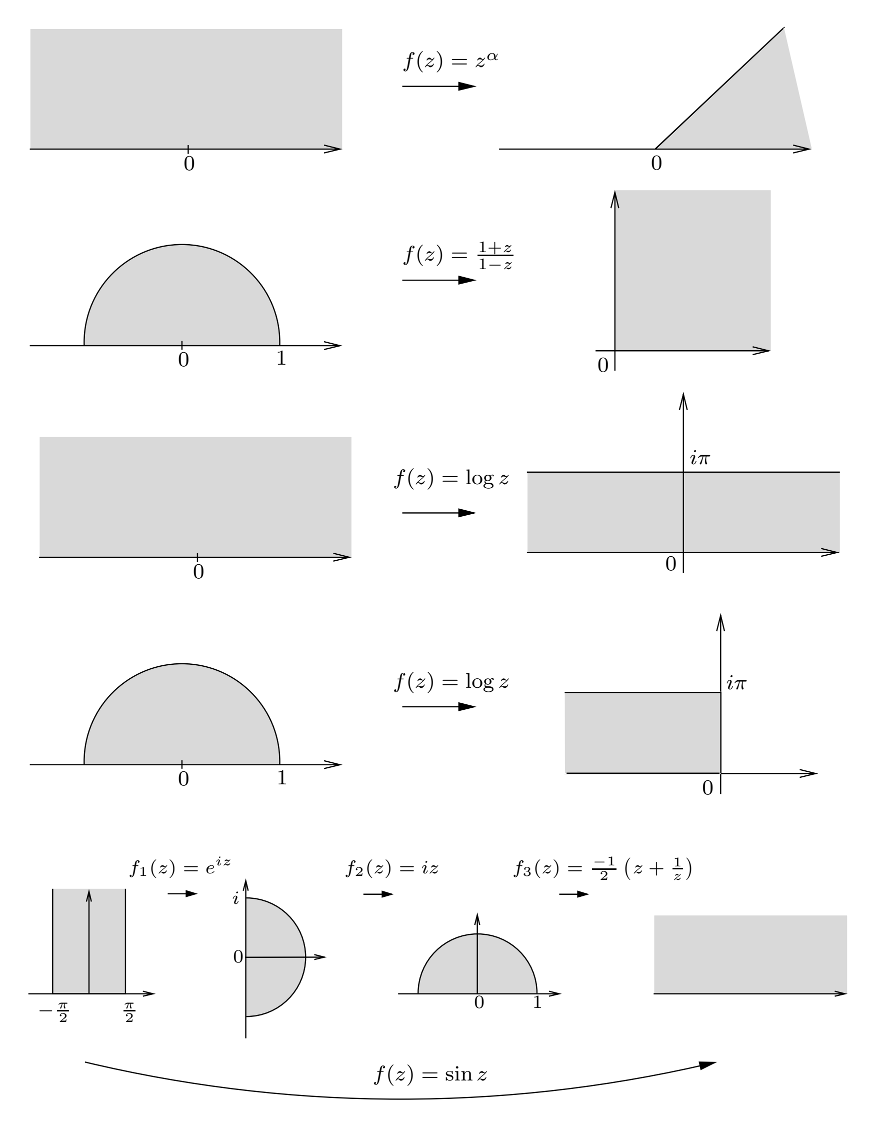


Figure 1. Explicit conformal maps

upper half-plane, where it arose in the solution of the steady-state heat equation. In these specific examples, explicit solutions were obtained in terms of convolutions with the Poisson kernels.

Our goal here is to connect the Dirichlet problem with the conformal maps discussed so far. We begin by providing a formula for a solution to the problem (1) in the special case where Ω is a strip. In fact, this exam-

ple was studied in Problem 3 of Chapter 5, Book I, where the problem was solved using the Fourier transform. Here, we recover this solution using only conformal mappings and the known solution in the disc.

The first important fact that we use is that the composition of a harmonic function with a holomorphic function is still harmonic.

Lemma 1.3 *Let V and U be open sets in $\mathbb{C}$ and $F : V \to U$ a holomorphic function. If $u : U \to \mathbb{C}$ is a harmonic function, then $u \circ F$ is harmonic on V.*

Proof. The thrust of the lemma is purely local, so we may assume that U is an open disc. We let G be a holomorphic function in U whose real part is u (such a G exists by Exercise 12 in Chapter 2, and is determined up to an additive constant). Let $H = G \circ F$ and note that $u \circ F$ is the real part of H. Hence $u \circ F$ is harmonic because H is holomorphic.

For an alternate (computational) proof of this lemma, see Exercise 6.

With this result in hand, we may now consider the problem (1) when Ω consists of the horizontal strip

$$\Omega = \{x + iy : x \in \mathbb{R},\ 0 < y < 1\},$$

whose boundary is the union of the two horizontal lines $\mathbb{R}$ and $i + \mathbb{R}$. We express the boundary data as two functions f_0 and f_1 defined on $\mathbb{R}$, and ask for a solution $u(x, y)$ in Ω of $\triangle u = 0$ that satisfies

$$u(x, 0) = f_0(x) \quad \text{and} \quad u(x, 1) = f_1(x).$$

We shall assume that f_0 and f_1 are continuous and vanish at infinity, that is, that $\lim_{|x|\to\infty} f_j(x) = 0$ for $j = 0, 1$.

The method we shall follow consists of relocating the problem from the strip to the unit disc via a conformal map. In the disc the solution $\tilde{u}$ is then expressed in terms of a convolution with the Poisson kernel. Finally, $\tilde{u}$ is moved back to the strip using the inverse of the previous conformal map, thereby giving our final answer to the problem.

To achieve our goal, we introduce the mappings $F : \mathbb{D} \to \Omega$ and $G : \Omega \to \mathbb{D}$, that are defined by

$$F(w) = \frac{1}{\pi} \log\left(i\frac{1-w}{1+w}\right) \quad \text{and} \quad G(z) = \frac{i - e^{\pi z}}{i + e^{\pi z}}.$$

These two functions, which are obtained from composing mappings from examples in the previous sections, are conformal and inverses to one

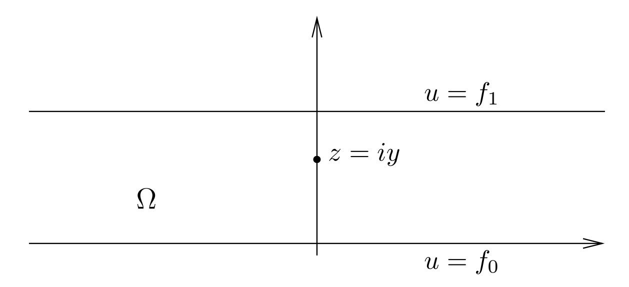


Figure 2. The Dirichlet problem in a strip

another. Tracing through the boundary behavior of F, we find that it maps the lower half-circle to the line $i + \mathbb{R}$, and the upper half-circle to $\mathbb{R}$. More precisely, as φ travels from $-\pi$ to 0, then $F(e^{i\varphi})$ goes from $i + \infty$ to $i - \infty$, and as φ travels from 0 to π, then $F(e^{i\varphi})$ goes from $-\infty$ to ∞ on the real line.

With the behavior of F on the circle in mind, we define

$$\tilde{f}_1(\varphi) = f_1(F(e^{i\varphi}) - i) \quad \text{whenever } -\pi < \varphi < 0,$$

and

$$\tilde{f}_0(\varphi) = f_0(F(e^{i\varphi})) \quad \text{whenever } 0 < \varphi < \pi.$$

Then, since f_0 and f_1 vanish at infinity, the function $\tilde{f}$ that is equal to $\tilde{f}_1$ on the lower semi-circle, $\tilde{f}_0$ on the upper semi-circle, and 0 at the points $\varphi = \pm\pi, 0$, is continuous on the whole circle. The solution to the Dirichlet problem in the unit disc with boundary data $\tilde{f}$ is given by the Poisson integral[3]

$$\begin{aligned}\tilde{u}(w) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta - \varphi)\tilde{f}(\varphi)\,d\varphi \\ &= \frac{1}{2\pi}\int_{-\pi}^{0} P_r(\theta - \varphi)\tilde{f}_1(\varphi)\,d\varphi + \frac{1}{2\pi}\int_{0}^{\pi} P_r(\theta - \varphi)\tilde{f}_0(\varphi)\,d\varphi,\end{aligned}$$

where $w = re^{i\theta}$, and

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}$$

[3]We refer the reader to Chapter 2 in Book I for a detailed discussion of the Dirichlet problem in the disc and the Poisson integral formula. Also, the Poisson integral formula is deduced in Exercise 12 of Chapter 2 and Problem 2 in Chapter 3 of this book.

is the Poisson kernel. Lemma 1.3 guarantees that the function u, defined by

$$u(z) = \tilde{u}(G(z)),$$

is harmonic in the strip. Moreover, our construction also insures that u has the correct boundary values.

A formula for u in terms of f_0 and f_1 is first obtained at the points $z = iy$ with $0 < y < 1$. The appropriate change of variables (see Exercise 7) shows that if $re^{i\theta} = G(iy)$, then

$$\frac{1}{2\pi}\int_0^{\pi} P_r(\theta - \varphi)\tilde{f}_0(\varphi)\,d\varphi = \frac{\sin \pi y}{2}\int_{-\infty}^{\infty} \frac{f_0(t)}{\cosh \pi t - \cos \pi y}\,dt.$$

A similar calculation also establishes

$$\frac{1}{2\pi}\int_0^{\pi} P_r(\theta - \varphi)\tilde{f}_1(\varphi)\,d\varphi = \frac{\sin \pi y}{2}\int_{-\infty}^{\infty} \frac{f_1(t)}{\cosh \pi t + \cos \pi y}\,dt.$$

Adding these last two integrals provides a formula for $u(0, y)$. In general, we recall from Exercise 13 in Chapter 5 of Book I, that a solution to the Dirichlet problem in the strip vanishing at infinity is unique. Consequently, a translation of the boundary condition by x results in a translation of the solution by x as well. We may therefore apply the same argument to $f_0(x + t)$ and $f_1(x + t)$ (with x fixed), and a final change of variables shows that

$$u(x, y) = \frac{\sin \pi y}{2}\left(\int_{-\infty}^{\infty} \frac{f_0(x - t)}{\cosh \pi t - \cos \pi y}\,dt + \int_{-\infty}^{\infty} \frac{f_1(x - t)}{\cosh \pi t + \cos \pi y}\,dt\right),$$

which gives a solution to the Dirichlet problem in the strip. In particular, we find that the solution is given in terms of convolutions with the functions f_0 and f_1. Also, note that at the mid-point of the strip ($y = 1/2$), the solution is given by integration with respect to the function $1/\cosh \pi t$; this function happens to be its own Fourier transform, as we saw in Example 3, Chapter 3.

Remarks about the Dirichlet problem

The example above leads us to envisage the solution of the more general Dirichlet problem for Ω (a suitable region), if we know a conformal map F from the disc $\mathbb{D}$ to Ω. That is, suppose we wish to solve (1), where f is an assigned continuous function and $\partial\Omega$ is the boundary of Ω. Assuming we have a conformal map F from $\mathbb{D}$ to Ω (that extends to a continuous

bijection of the boundary of the disc to the boundary of Ω), then $\tilde{f} = f \circ F$ is defined on the circle, and we can solve the Dirichlet problem for the disc with boundary data $\tilde{f}$. The solution is given by the Poisson integral formula

$$\tilde{u}(re^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} P_r(\theta - \varphi) \tilde{f}(e^{i\varphi})\, d\varphi,$$

where P_r is the Poisson kernel. Then, one can expect that the solution of the original problem is given by $u = \tilde{u} \circ F^{-1}$.

Success with this approach requires that we are able to resolve affirmatively two questions:

- Does there exist a conformal map $\Phi = F^{-1}$ from Ω to $\mathbb{D}$?
- If so, does this map extend to a continuous bijection from the boundary of Ω to the boundary of $\mathbb{D}$?

The first question, that of existence, is settled by the Riemann mapping theorem, which we prove in the next section. It is completely general (assuming only that Ω is a proper subset of $\mathbb{C}$ that is simply connected), and necessitates no regularity of the boundary of Ω. A positive answer to the second question requires some regularity of $\partial\Omega$. A particular case, when Ω is the interior of a polygon, is treated below in Section 4.3. (See Exercise 18 and Problem 6 for more general assertions.)

It is interesting to note that in Riemann's original approach to the mapping problem, the chain of implications was reversed: his idea was that the existence of the conformal map Φ from Ω to $\mathbb{D}$ is a consequence of the solvability of the Dirichlet problem in Ω. He argued as follows. Suppose we wish to find such a Φ, with the property that a given point $z_0 \in \Omega$ is mapped to 0. Then Φ must be of the form

$$\Phi(z) = (z - z_0)G(z),$$

where G is holomorphic and non-vanishing in Ω. Hence we can take

$$\Phi(z) = (z - z_0)e^{H(z)},$$

for suitable H. Now if $u(z)$ is the harmonic function given by $u = \mathrm{Re}(H)$, then the fact that $|\Phi(z)| = 1$ on $\partial\Omega$ means that u must satisfy the boundary condition $u(z) = \log(1/|z - z_0|)$ for $z \in \partial\Omega$. So if we can find such a solution u of the Dirichlet problem,[4] we can construct H, and from this the mapping function Φ.

---

[4]The harmonic function $u(z)$ is also known as the Green's function with source z_0 for the region Ω.

However, there are several shortcomings to this method. First, one has to verify that Φ is a bijection. In addition, to succeed, this method requires some regularity of the boundary of Ω. Moreover, one is still faced with the question of solving the Dirichlet problem for Ω. At this stage Riemann proposed using the "Dirichlet principle." But applying this idea involves difficulties that must be overcome.[5]

Nevertheless, using different methods, one can prove the existence of the mapping in the general case. This approach is carried out below in Section 3.

2 The Schwarz lemma; automorphisms of the disc and upper half-plane

The statement and proof of the Schwarz lemma are both simple, but the applications of this result are far-reaching. We recall that a rotation is a map of the form $z \mapsto cz$ with $|c| = 1$, namely $c = e^{i\theta}$, where $\theta \in \mathbb{R}$ is called the angle of rotation and is well-defined up to an integer multiple of 2π.

Lemma 2.1 *Let $f : \mathbb{D} \to \mathbb{D}$ be holomorphic with $f(0) = 0$. Then*

(i) $|f(z)| \leq |z|$ *for all* $z \in \mathbb{D}$.

(ii) *If for some* $z_0 \neq 0$ *we have* $|f(z_0)| = |z_0|$, *then* f *is a rotation.*

(iii) $|f'(0)| \leq 1$, *and if equality holds, then* f *is a rotation.*

Proof. We first expand f in a power series centered at 0 and convergent in all of $\mathbb{D}$

$$f(z) = a_0 + a_1 z + a_2 z^2 + \cdots .$$

Since $f(0) = 0$ we have $a_0 = 0$, and therefore $f(z)/z$ is holomorphic in $\mathbb{D}$ (since it has a removable singularity at 0). If $|z| = r < 1$, then since $|f(z)| \leq 1$ we have

$$\left| \frac{f(z)}{z} \right| \leq \frac{1}{r},$$

and by the maximum modulus principle, we can conclude that this is true whenever $|z| \leq r$. Letting $r \to 1$ gives the first result.

For (ii), we see that $f(z)/z$ attains its maximum in the interior of $\mathbb{D}$ and must therefore be constant, say $f(z) = cz$. Evaluating this expression

---

[5] An implementation of Dirichlet's principle in the present two-dimensional situation is taken up in Book III.

at z_0 and taking absolute values, we find that $|c| = 1$. Therefore, there exists $\theta \in \mathbb{R}$ such that $c = e^{i\theta}$, and that explains why f is a rotation.

Finally, observe that if $g(z) = f(z)/z$, then $|g(z)| \leq 1$ throughout $\mathbb{D}$, and moreover

$$g(0) = \lim_{z \to 0} \frac{f(z) - f(0)}{z} = f'(0).$$

Hence, if $|f'(0)| = 1$, then $|g(0)| = 1$, and by the maximum principle g is constant, which implies $f(z) = cz$ with $|c| = 1$.

Our first application of this lemma is to the determination of the automorphisms of the disc.

2.1 Automorphisms of the disc

A conformal map from an open set Ω to *itself* is called an **automorphism** of Ω. The set of all automorphisms of Ω is denoted by $\mathrm{Aut}(\Omega)$, and carries the structure of a group. The group operation is composition of maps, the identity element is the map $z \mapsto z$, and the inverses are simply the inverse functions. It is clear that if f and g are automorphisms of Ω, then $f \circ g$ is also an automorphism, and in fact, its inverse is given by

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

As mentioned above, the identity map is always an automorphism. We can give other more interesting automorphisms of the unit disc. Obviously, any rotation by an angle $\theta \in \mathbb{R}$, that is, $r_\theta : z \mapsto e^{i\theta} z$, is an automorphism of the unit disc whose inverse is the rotation by the angle $-\theta$, that is, $r_{-\theta} : z \mapsto e^{-i\theta} z$. More interesting, are the automorphisms of the form

$$\psi_\alpha(z) = \frac{\alpha - z}{1 - \overline{\alpha} z}, \qquad \text{where } \alpha \in \mathbb{C} \text{ with } |\alpha| < 1.$$

These mappings, which where introduced in Exercise 7 of Chapter 1, appear in a number of problems in complex analysis because of their many useful properties. The proof that they are automorphisms of $\mathbb{D}$ is quite simple. First, observe that since $|\alpha| < 1$, the map ψ_α is holomorphic in the unit disc. If $|z| = 1$ then $z = e^{i\theta}$ and

$$\psi_\alpha(e^{i\theta}) = \frac{\alpha - e^{i\theta}}{e^{i\theta}(e^{-i\theta} - \overline{\alpha})} = e^{-i\theta} \frac{w}{\overline{w}},$$

where $w = \alpha - e^{i\theta}$, therefore $|\psi_\alpha(z)| = 1$. By the maximum modulus principle, we conclude that $|\psi_\alpha(z)| < 1$ for all $z \in \mathbb{D}$. Finally we make

the following very simple observation:

$$\begin{aligned}(\psi_\alpha \circ \psi_\alpha)(z) &= \frac{\alpha - \frac{\alpha - z}{1-\overline{\alpha}z}}{1 - \overline{\alpha}\frac{\alpha - z}{1-\overline{\alpha}z}} \\ &= \frac{\alpha - |\alpha|^2 z - \alpha + z}{1 - \overline{\alpha}z - |\alpha|^2 + \overline{\alpha}z} \\ &= \frac{(1-|\alpha|^2)z}{1-|\alpha|^2} \\ &= z,\end{aligned}$$

from which we conclude that ψ_α is its own inverse! Another important property of ψ_α is that it vanishes at $z = \alpha$; moreover it interchanges 0 and α, namely

$$\psi_\alpha(0) = \alpha \quad \text{and} \quad \psi_\alpha(\alpha) = 0.$$

The next theorem says that the rotations combined with the maps ψ_α exhaust all the automorphisms of the disc.

Theorem 2.2 *If f is an automorphism of the disc, then there exist $\theta \in \mathbb{R}$ and $\alpha \in \mathbb{D}$ such that*

$$f(z) = e^{i\theta}\frac{\alpha - z}{1-\overline{\alpha}z}.$$

Proof. Since f is an automorphism of the disc, there exists a unique complex number $\alpha \in \mathbb{D}$ such that $f(\alpha) = 0$. Now we consider the automorphism g defined by $g = f \circ \psi_\alpha$. Then $g(0) = 0$, and the Schwarz lemma gives

$$|g(z)| \le |z| \quad \text{for all } z \in \mathbb{D}. \tag{2}$$

Moreover, $g^{-1}(0) = 0$, so applying the Schwarz lemma to g^{-1}, we find that

$$|g^{-1}(w)| \le |w| \quad \text{for all } w \in \mathbb{D}.$$

Using this last inequality for $w = g(z)$ for each $z \in \mathbb{D}$ gives

$$|z| \le |g(z)| \quad \text{for all } z \in \mathbb{D}. \tag{3}$$

Combining (2) and (3) we find that $|g(z)| = |z|$ for all $z \in \mathbb{D}$, and by the Schwarz lemma we conclude that $g(z) = e^{i\theta}z$ for some $\theta \in \mathbb{R}$. Replacing z by $\psi_\alpha(z)$ and using the fact that $(\psi_\alpha \circ \psi_\alpha)(z) = z$, we deduce that $f(z) = e^{i\theta}\psi_\alpha(z)$, as claimed.

Setting $\alpha = 0$ in the theorem yields the following result.

Corollary 2.3 *The only automorphisms of the unit disc that fix the origin are the rotations.*

Note that by the use of the mappings ψ_α, we can see that the group of automorphisms of the disc acts **transitively**, in the sense that given any pair of points α and β in the disc, there is an automorphism ψ mapping α to β. One such ψ is given by $\psi = \psi_\beta \circ \psi_\alpha$.

The explicit formulas for the automorphisms of $\mathbb{D}$ give a good description of the group $\text{Aut}(\mathbb{D})$. In fact, this group of automorphisms is "almost" isomorphic to a group of 2×2 matrices with complex entries often denoted by $\text{SU}(1,1)$. This group consists of all 2×2 matrices that preserve the hermitian form on $\mathbb{C}^2 \times \mathbb{C}^2$ defined by

$$\langle Z, W \rangle = z_1 \overline{w}_1 - z_2 \overline{w}_2,$$

where $Z = (z_1, z_2)$ and $W = (w_1, w_2)$. For more information about this subject, we refer the reader to Problem 4.

2.2 Automorphisms of the upper half-plane

Our knowledge of the automorphisms of $\mathbb{D}$ together with the conformal map $F : \mathbb{H} \to \mathbb{D}$ found in Section 1.1 allow us to determine the group of automorphisms of $\mathbb{H}$ which we denote by $\text{Aut}(\mathbb{H})$.

Consider the map

$$\Gamma : \text{Aut}(\mathbb{D}) \to \text{Aut}(\mathbb{H})$$

given by "conjugation by F":

$$\Gamma(\varphi) = F^{-1} \circ \varphi \circ F.$$

It is clear that $\Gamma(\varphi)$ is an automorphism of $\mathbb{H}$ whenever φ is an automorphism of $\mathbb{D}$, and Γ is a bijection whose inverse is given by $\Gamma^{-1}(\psi) = F \circ \psi \circ F^{-1}$. In fact, we prove more, namely that Γ preserves the operations on the corresponding groups of automorphisms. Indeed, suppose that $\varphi_1, \varphi_2 \in \text{Aut}(\mathbb{D})$. Since $F \circ F^{-1}$ is the identity on $\mathbb{D}$ we find that

$$\begin{aligned}
\Gamma(\varphi_1 \circ \varphi_2) &= F^{-1} \circ \varphi_1 \circ \varphi_2 \circ F \\
&= F^{-1} \circ \varphi_1 \circ F \circ F^{-1} \circ \varphi_2 \circ F \\
&= \Gamma(\varphi_1) \circ \Gamma(\varphi_2).
\end{aligned}$$

The conclusion is that the two groups $\text{Aut}(\mathbb{D})$ and $\text{Aut}(\mathbb{H})$ are the same, since Γ defines an isomorphism between them. We are still left with the

task of giving a description of elements of $\text{Aut}(\mathbb{H})$. A series of calculations, which consist of pulling back the automorphisms of the disc to the upper half-plane via F, can be used to verify that $\text{Aut}(\mathbb{H})$ consists of all maps

$$z \mapsto \frac{az+b}{cz+d},$$

where a, b, c, and d are real numbers with $ad - bc = 1$. Again, a matrix group is lurking in the background. Let $\text{SL}_2(\mathbb{R})$ denote the group of all 2×2 matrices with real entries and determinant 1, namely

$$\text{SL}_2(\mathbb{R}) = \left\{ M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \quad a, b, c, d \in \mathbb{R} \text{ and } \det(M) = ad - bc = 1 \right\}.$$

This group is called the **special linear group**.

Given a matrix $M \in \text{SL}_2(\mathbb{R})$ we define the mapping f_M by

$$f_M(z) = \frac{az+b}{cz+d}.$$

Theorem 2.4 *Every automorphism of* $\mathbb{H}$ *takes the form* f_M *for some* $M \in \text{SL}_2(\mathbb{R})$. *Conversely, every map of this form is an automorphism of* $\mathbb{H}$.

The proof consists of a sequence of steps. For brevity, we denote the group $\text{SL}_2(\mathbb{R})$ by $\mathcal{G}$.

Step 1. If $M \in \mathcal{G}$, then f_M maps $\mathbb{H}$ to itself. This is clear from the observation that

$$\text{(4)} \quad \text{Im}(f_M(z)) = \frac{(ad-bc)\text{Im}(z)}{|cz+d|^2} = \frac{\text{Im}(z)}{|cz+d|^2} > 0 \quad \text{whenever } z \in \mathbb{H}.$$

Step 2. If M and M' are two matrices in $\mathcal{G}$, then $f_M \circ f_{M'} = f_{MM'}$. This follows from a straightforward calculation, which we omit. As a consequence, we can prove the first half of the theorem. Each f_M is an automorphism because it has a holomorphic inverse $(f_M)^{-1}$, which is simply $f_{M^{-1}}$. Indeed, if I is the identity matrix, then

$$(f_M \circ f_{M^{-1}})(z) = f_{MM^{-1}}(z) = f_I(z) = z.$$

Step 3. Given any two points z and w in $\mathbb{H}$, there exists $M \in \mathcal{G}$ such that $f_M(z) = w$, and therefore $\mathcal{G}$ acts transitively on $\mathbb{H}$. To prove this,

it suffices to show that we can map any $z \in \mathbb{H}$ to i. Setting $d = 0$ in equation (4) above gives

$$\mathrm{Im}(f_M(z)) = \frac{\mathrm{Im}(z)}{|cz|^2}$$

and we may choose a real number c so that $\mathrm{Im}(f_M(z)) = 1$. Next we choose the matrix

$$M_1 = \begin{pmatrix} 0 & -c^{-1} \\ c & 0 \end{pmatrix}$$

so that $f_{M_1}(z)$ has imaginary part equal to 1. Then we translate by a matrix of the form

$$M_2 = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \qquad \text{with } b \in \mathbb{R},$$

to bring $f_{M_1}(z)$ to i. Finally, the map f_M with $M = M_2 M_1$ takes z to i.

Step 4. If θ is real, then the matrix

$$M_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

belongs to $\mathcal{G}$, and if $F : \mathbb{H} \to \mathbb{D}$ denotes the standard conformal map, then $F \circ f_{M_\theta} \circ F^{-1}$ corresponds to the rotation of angle -2θ in the disc. This follows from the fact that $F \circ f_{M_\theta} = e^{-2i\theta} F(z)$, which is easily verified.

Step 5. We can now complete the proof of the theorem. We suppose f is an automorphism of $\mathbb{H}$ with $f(\beta) = i$, and consider a matrix $N \in \mathcal{G}$ such that $f_N(i) = \beta$. Then $g = f \circ f_N$ satisfies $g(i) = i$, and therefore $F \circ g \circ F^{-1}$ is an automorphism of the disc that fixes the origin. So $F \circ g \circ F^{-1}$ is a rotation, and by Step 4 there exists $\theta \in \mathbb{R}$ such that

$$F \circ g \circ F^{-1} = F \circ f_{M_\theta} \circ F^{-1}.$$

Hence $g = f_{M_\theta}$, and we conclude that $f = f_{M_\theta N^{-1}}$ which is of the desired form.

A final observation is that the group $\mathrm{Aut}(\mathbb{H})$ is not quite isomorphic with $\mathrm{SL}_2(\mathbb{R})$. The reason for this is because the two matrices M and $-M$ give rise to the same function $f_M = f_{-M}$. Therefore, if we identify the two matrices M and $-M$, then we obtain a new group $\mathrm{PSL}_2(\mathbb{R})$ called the **projective special linear group**; this group is isomorphic with $\mathrm{Aut}(\mathbb{H})$.

3 The Riemann mapping theorem

3.1 Necessary conditions and statement of the theorem

We now come to the promised cornerstone of this chapter. The basic problem is to determine conditions on an open set Ω that guarantee the existence of a conformal map $F : \Omega \to \mathbb{D}$.

A series of simple observations allow us to find necessary conditions on Ω. First, if $\Omega = \mathbb{C}$ there can be no conformal map $F : \Omega \to \mathbb{D}$, since by Liouville's theorem F would have to be a constant. Therefore, a necessary condition is to assume that $\Omega \neq \mathbb{C}$. Since $\mathbb{D}$ is connected, we must also impose the requirement that Ω be connected. There is still one more condition that is forced upon us: since $\mathbb{D}$ is simply connected, the same must be true of Ω (see Exercise 3). It is remarkable that these conditions on Ω are also sufficient to guarantee the existence of a biholomorpism from Ω to $\mathbb{D}$.

For brevity, we shall call a subset Ω of $\mathbb{C}$ **proper** if it is non-empty and not the whole of $\mathbb{C}$.

Theorem 3.1 (Riemann mapping theorem) *Suppose Ω is proper and simply connected. If $z_0 \in \Omega$, then there exists a unique conformal map $F : \Omega \to \mathbb{D}$ such that*

$$F(z_0) = 0 \quad \text{and} \quad F'(z_0) > 0.$$

Corollary 3.2 *Any two proper simply connected open subsets in $\mathbb{C}$ are conformally equivalent.*

Clearly, the corollary follows from the theorem, since we can use as an intermediate step the unit disc. Also, the uniqueness statement in the theorem is straightforward, since if F and G are conformal maps from Ω to $\mathbb{D}$ that satisfy these two conditions, then $H = F \circ G^{-1}$ is an automorphism of the disc that fixes the origin. Therefore $H(z) = e^{i\theta}z$, and since $H'(0) > 0$, we must have $e^{i\theta} = 1$, from which we conclude that $F = G$.

The rest of this section is devoted to the proof of the existence of the conformal map F. The idea of the proof is as follows. We consider all injective holomorphic maps $f : \Omega \to \mathbb{D}$ with $f(z_0) = 0$. From these we wish to choose an f so that its image fills out all of $\mathbb{D}$, and this can be achieved by making $f'(z_0)$ as large as possible. In doing this, we shall need to be able to extract f as a limit from a given sequence of functions. We turn to this point first.

3.2 Montel's theorem

Let Ω be an open subset of $\mathbb{C}$. A family $\mathcal{F}$ of holomorphic functions on Ω is said to be **normal** if every sequence in $\mathcal{F}$ has a subsequence that converges uniformly on every compact subset of Ω (the limit need not be in $\mathcal{F}$).

The proof that a family of functions is normal is, in practice, the consequence of two related properties, uniform boundedness and equicontinuity. These we shall now define.

The family $\mathcal{F}$ is said to be **uniformly bounded on compact subsets of** Ω if for each compact set $K \subset \Omega$ there exists $B > 0$, such that

$$|f(z)| \leq B \quad \text{for all } z \in K \text{ and } f \in \mathcal{F}.$$

Also, the family $\mathcal{F}$ is **equicontinuous** on a compact set K if for every $\epsilon > 0$ there exists $\delta > 0$ such that whenever $z, w \in K$ and $|z - w| < \delta$, then

$$|f(z) - f(w)| < \epsilon \quad \text{for all } f \in \mathcal{F}.$$

Equicontinuity is a strong condition, which requires uniform continuity, uniformly in the family. For instance, any family of differentiable functions on $[0, 1]$ whose derivatives are uniformly bounded is equicontinuous. This follows directly from the mean value theorem. On the other hand, note that the family $\{f_n\}$ on $[0, 1]$ given by $f_n(x) = x^n$ is *not* equicontinuous since for any fixed $0 < x_0 < 1$ we have $|f_n(1) - f_n(x_0)| \to 1$ as n tends to infinity.

The theorem that follows puts together these new concepts and is an important ingredient in the proof of the Riemann mapping theorem.

Theorem 3.3 *Suppose $\mathcal{F}$ is a family of holomorphic functions on Ω that is uniformly bounded on compact subsets of Ω. Then:*

(i) *$\mathcal{F}$ is equicontinuous on every compact subset of Ω.*

(ii) *$\mathcal{F}$ is a normal family.*

The theorem really consists of two separate parts. The first part says that $\mathcal{F}$ is equicontinuous under the assumption that $\mathcal{F}$ is a family of *holomorphic* functions that is uniformly bounded on compact subsets of Ω. The proof follows from an application of the Cauchy integral formula and hence relies on the fact that $\mathcal{F}$ consists of holomorphic functions. This conclusion is in sharp contrast with the real situation as illustrated by the family of functions given by $f_n(x) = \sin(nx)$ on $(0, 1)$, which is

uniformly bounded. However, this family is not equicontinuous and has no convergent subsequence on any compact subinterval of $(0,1)$.

The second part of the theorem is not complex-analytic in nature. Indeed, the fact that $\mathcal{F}$ is a normal family follows from assuming only that $\mathcal{F}$ is uniformly bounded and equicontinuous on compact subsets of Ω. This result is sometimes known as the Arzela-Ascoli theorem and its proof consists primarily of a diagonalization argument.

We are required to prove convergence on arbitrary compact subsets of Ω, therefore it is useful to introduce the following notion. A sequence $\{K_\ell\}_{\ell=1}^{\infty}$ of compact subsets of Ω is called an **exhaustion** if

(a) K_ℓ is contained in the interior of $K_{\ell+1}$ for all $\ell = 1, 2, \ldots$.

(b) Any compact set $K \subset \Omega$ is contained in K_ℓ for some ℓ. In particular
$$\Omega = \bigcup_{\ell=1}^{\infty} K_\ell.$$

Lemma 3.4 *Any open set Ω in the complex plane has an exhaustion.*

Proof. If Ω is bounded, we let K_ℓ denote the set of all points in Ω at distance $\geq 1/\ell$ from the boundary of Ω. If Ω is not bounded, let K_ℓ denote the same set as above except that we also require $|z| \leq \ell$ for all $z \in K_\ell$.

We may now begin the proof of Montel's theorem. Let K be a compact subset of Ω and choose $r > 0$ so small that $D_{3r}(z)$ is contained in Ω for all $z \in K$. It suffices to choose r so that $3r$ is less than the distance from K to the boundary of Ω. Let $z, w \in K$ with $|z - w| < r$, and let γ denote the boundary circle of the disc $D_{2r}(w)$. Then, by Cauchy's integral formula, we have

$$f(z) - f(w) = \frac{1}{2\pi i} \int_\gamma f(\zeta) \left[\frac{1}{\zeta - z} - \frac{1}{\zeta - w} \right] d\zeta.$$

Observe that

$$\left| \frac{1}{\zeta - z} - \frac{1}{\zeta - w} \right| = \frac{|z - w|}{|\zeta - z|\,|\zeta - w|} \leq \frac{|z - w|}{r^2}$$

since $\zeta \in \gamma$ and $|z - w| < r$. Therefore

$$|f(z) - f(w)| \leq \frac{1}{2\pi} \frac{2\pi r}{r^2} B|z - w|,$$

where B denotes the uniform bound for the family $\mathcal{F}$ in the compact set consisting of all points in Ω at a distance $\leq 2r$ from K. Therefore $|f(z) - f(w)| < C|z - w|$, and this estimate is true for all $z, w \in K$ with $|z - w| < r$ and $f \in \mathcal{F}$; thus this family is equicontinuous, as was to be shown.

To prove the second part of the theorem, we argue as follows. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{F}$ and K a compact subset of Ω. Choose a sequence of points $\{w_j\}_{j=1}^{\infty}$ that is dense in Ω. Since $\{f_n\}$ is uniformly bounded, there exists a subsequence $\{f_{n,1}\} = \{f_{1,1}, f_{2,1}, f_{3,1}, \ldots\}$ of $\{f_n\}$ such that $f_{n,1}(w_1)$ converges.

From $\{f_{n,1}\}$ we can extract a subsequence $\{f_{n,2}\} = \{f_{1,2}, f_{2,2}, f_{3,2}, \ldots\}$ so that $f_{n,2}(w_2)$ converges. We may continue this process, and extract a subsequence $\{f_{n,j}\}$ of $\{f_{n,j-1}\}$ such that $f_{n,j}(w_j)$ converges.

Finally, let $g_n = f_{n,n}$ and consider the diagonal subsequence $\{g_n\}$. By construction, $g_n(w_j)$ converges for each j, and we claim that equicontinuity implies that g_n converges uniformly on K. Given $\epsilon > 0$, choose δ as in the definition of equicontinuity, and note that for some J, the set K is contained in the union of the discs $D_\delta(w_1), \ldots, D_\delta(w_J)$. Pick N so large that if $n, m > N$, then

$$|g_m(w_j) - g_n(w_j)| < \epsilon \quad \text{for all } j = 1, \ldots, J.$$

So if $z \in K$, then $z \in D_\delta(w_j)$ for some $1 \leq j \leq J$. Therefore,

$$|g_n(z) - g_m(z)| \leq |g_n(z) - g_n(w_j)| + |g_n(w_j) - g_m(w_j)| +$$

$$+ |g_m(w_j) - g_m(z)| < 3\epsilon$$

whenever $n, m > N$. Hence $\{g_n\}$ converges uniformly on K.

Finally, we need one more diagonalization argument to obtain a subsequence that converges uniformly on *every* compact subset of Ω. Let $K_1 \subset K_2 \subset \cdots \subset K_\ell \subset \cdots$ be an exhaustion of Ω, and suppose $\{g_{n,1}\}$ is a subsequence of the original sequence $\{f_n\}$ that converges uniformly on K_1. Extract from $\{g_{n,1}\}$ a subsequence $\{g_{n,2}\}$ that converges uniformly on K_2, and so on. Then, $\{g_{n,n}\}$ is a subsequence of $\{f_n\}$ that converges uniformly on every K_ℓ and since the K_ℓ exhaust Ω, the sequence $\{g_{n,n}\}$ converges uniformly on any compact subset of Ω, as was to be shown.

We need one further result before we can give the proof of the Riemann mapping theorem.

Proposition 3.5 *If Ω is a connected open subset of $\mathbb{C}$ and $\{f_n\}$ a sequence of injective holomorphic functions on Ω that converges uniformly*

on every compact subset of Ω to a holomorphic function f, then f is either injective or constant.

Proof. We argue by contradiction and suppose that f is not injective, so there exist distinct complex numbers z_1 and z_2 in Ω such that $f(z_1) = f(z_2)$. Define a new sequence by $g_n(z) = f_n(z) - f_n(z_1)$, so that g_n has no other zero besides z_1, and the sequence $\{g_n\}$ converges uniformly on compact subsets of Ω to $g(z) = f(z) - f(z_1)$. If g is not identically zero, then z_2 is an isolated zero for g (because Ω is connected); therefore

$$1 = \frac{1}{2\pi i} \int_\gamma \frac{g'(\zeta)}{g(\zeta)}\, d\zeta,$$

where γ is a small circle centered at z_2 chosen so that g does not vanish on γ or at any point of its interior besides z_2. Therefore, $1/g_n$ converges uniformly to $1/g$ on γ, and since $g'_n \to g'$ uniformly on γ we have

$$\frac{1}{2\pi i} \int_\gamma \frac{g'_n(\zeta)}{g_n(\zeta)}\, d\zeta \to \frac{1}{2\pi i} \int_\gamma \frac{g'(\zeta)}{g(\zeta)}\, d\zeta.$$

But this is a contradiction since g_n has no zeros inside γ, and hence

$$\frac{1}{2\pi i} \int_\gamma \frac{g'_n(\zeta)}{g_n(\zeta)}\, d\zeta = 0 \qquad \text{for all } n\,.$$

3.3 Proof of the Riemann mapping theorem

Once we have established the technical results above, the rest of the proof of the Riemann mapping theorem is very elegant. It consists of three steps, which we isolate.

Step 1. Suppose that Ω is a simply connected proper open subset of $\mathbb{C}$. We claim that Ω is conformally equivalent to an open subset of the unit disc that contains the origin. Indeed, choose a complex number α that does not belong to Ω, (recall that Ω is proper), and observe that $z - \alpha$ never vanishes on the simply connected set Ω. Therefore, we can define a holomorphic function

$$f(z) = \log(z - \alpha)$$

with the desired properties of the logarithm. As a consequence one has, $e^{f(z)} = z - \alpha$, which proves in particular that f is injective. Pick a point $w \in \Omega$, and observe that

$$f(z) \neq f(w) + 2\pi i \qquad \text{for all } z \in \Omega$$

for otherwise, we exponentiate this relation to find that $z = w$, hence $f(z) = f(w)$, a contradiction. In fact, we claim that $f(z)$ stays strictly away from $f(w) + 2\pi i$, in the sense that there exists a disc centered at $f(w) + 2\pi i$ that contains no points of the image $f(\Omega)$. Otherwise, there exists a sequence $\{z_n\}$ in Ω such that $f(z_n) \to f(w) + 2\pi i$. We exponentiate this relation, and, since the exponential function is continuous, we must have $z_n \to w$. But this implies $f(z_n) \to f(w)$, which is a contradiction. Finally, consider the map

$$F(z) = \frac{1}{f(z) - (f(w) + 2\pi i)}.$$

Since f is injective, so is F, hence $F : \Omega \to F(\Omega)$ is a conformal map. Moreover, by our analysis, $F(\Omega)$ is bounded. We may therefore translate and rescale the function F in order to obtain a conformal map from Ω to an open subset of $\mathbb{D}$ that contains the origin.

Step 2. By the first step, we may assume that Ω is an open subset of $\mathbb{D}$ with $0 \in \Omega$. Consider the family $\mathcal{F}$ of all injective holomorphic functions on Ω that map into the unit disc and fix the origin:

$$\mathcal{F} = \{f : \Omega \to \mathbb{D} \text{ holomorphic, injective and } f(0) = 0\}.$$

First, note that $\mathcal{F}$ is non-empty since it contains the identity. Also, this family is uniformly bounded by construction, since all functions are required to map into the unit disc.

Now, we turn to the question of finding a function $f \in \mathcal{F}$ that maximizes $|f'(0)|$. First, observe that the quantities $|f'(0)|$ are uniformly bounded as f ranges in $\mathcal{F}$. This follows from the Cauchy inequality (Corollary 4.3 in Chapter 2) for f' applied to a small disc centered at the origin.

Next, we let

$$s = \sup_{f \in \mathcal{F}} |f'(0)|,$$

and we choose a sequence $\{f_n\} \subset \mathcal{F}$ such that $|f_n'(0)| \to s$ as $n \to \infty$. By Montel's theorem (Theorem 3.3), this sequence has a subsequence that converges uniformly on compact sets to a holomorphic function f on Ω. Since $s \geq 1$ (because $z \mapsto z$ belongs to $\mathcal{F}$), f is non-constant, hence injective, by Proposition 3.5. Also, by continuity we have $|f(z)| \leq 1$ for all $z \in \Omega$ and from the maximum modulus principle we see that $|f(z)| < 1$. Since we clearly have $f(0) = 0$, we conclude that $f \in \mathcal{F}$ with $|f'(0)| = s$.

Step 3. In this last step, we demonstrate that f is a conformal map from Ω to $\mathbb{D}$. Since f is already injective, it suffices to prove that f is also surjective. If this were not true, we could construct a function in $\mathcal{F}$ with derivative at 0 greater than s. Indeed, suppose there exists $\alpha \in \mathbb{D}$ such that $f(z) \neq \alpha$, and consider the automorphism ψ_α of the disc that interchanges 0 and α, namely

$$\psi_\alpha(z) = \frac{\alpha - z}{1 - \overline{\alpha}z}.$$

Since Ω is simply connected, so is $U = (\psi_\alpha \circ f)(\Omega)$, and moreover, U does not contain the origin. It is therefore possible to define a square root function on U by

$$g(w) = e^{\frac{1}{2}\log w}.$$

Next, consider the function

$$F = \psi_{g(\alpha)} \circ g \circ \psi_\alpha \circ f.$$

We claim that $F \in \mathcal{F}$. Clearly F is holomorphic and it maps 0 to 0. Also F maps into the unit disc since this is true of each of the functions in the composition. Finally, F is injective. This is clearly true for the automorphisms ψ_α and $\psi_{g(\alpha)}$; it is also true for the square root g and the function f, since the latter is injective by assumption. If h denotes the square function $h(w) = w^2$, then we must have

$$f = \psi_\alpha^{-1} \circ h \circ \psi_{g(\alpha)}^{-1} \circ F = \Phi \circ F.$$

But Φ maps $\mathbb{D}$ into $\mathbb{D}$ with $\Phi(0) = 0$, and is not injective because F is and h is not. By the last part of the Schwarz lemma, we conclude that $|\Phi'(0)| < 1$. The proof is complete once we observe that

$$f'(0) = \Phi'(0)F'(0),$$

and thus

$$|f'(0)| < |F'(0)|,$$

contradicting the maximality of $|f'(0)|$ in $\mathcal{F}$.

Finally, we multiply f by a complex number of absolute value 1 so that $f'(0) > 0$, which ends the proof.

For a variant of this proof, see Problem 7.

Remark. It is worthwhile to point out that the only places where the hypothesis of simple-connectivity entered in the proof were in the uses of the logarithm and the square root. Thus it would have sufficed to have assumed (in addition to the hypothesis that Ω is proper) that Ω is **holomorphically simply connected** in the sense that for any holomorphic function f in Ω and any closed curve γ in Ω, we have $\int_\gamma f(z)\,dz = 0$. Further discussion of this point, and various equivalent properties of simple-connectivity, are given in Appendix B.

4 Conformal mappings onto polygons

The Riemann mapping theorem guarantees the existence of a conformal map from any proper, simply connected open set to the disc, or equivalently to the upper half-plane, but this theorem gives little insight as to the exact form of this map. In Section 1 we gave various explicit formulas in the case of regions that have symmetries, but it is of course unreasonable to ask for an explicit formula in the general case. There is, however, another class of open sets for which there are nice formulas, namely the polygons. Our aim in this last section is to give a proof of the Schwarz-Christoffel formula, which describes the nature of conformal maps from the disc (or upper half-plane) to polygons.

4.1 Some examples

We begin by studying some motivating examples. The first two correspond to easy (but infinite and degenerate) cases.

Example 1. First, we investigate the conformal map from the upper half-plane to the sector $\{z : 0 < \arg z < \alpha\pi\}$, with $0 < \alpha < 2$, given in Section 1 by $f(z) = z^\alpha$. Anticipating the Schwarz-Christoffel formula below, we write

$$z^\alpha = f(z) = \int_0^z f'(\zeta)\,d\zeta = \alpha \int_0^z \zeta^{-\beta}\,d\zeta$$

with $\alpha + \beta = 1$, and where the integral is taken along any path in the upper half-plane. In fact, by continuity and Cauchy's theorem, we may take the path of integration to lie in the closure of the upper half-plane. Although the behavior of f follows immediately from the original definition, we study it in terms of the integral expression above, since this provides insight for the general case treated later.

Note first that $\zeta^{-\beta}$ is integrable near 0 since $\beta < 1$, therefore $f(0) = 0$. Observe that when z is real and positive $(z = x)$, then $f'(x) = \alpha x^{\alpha-1}$ is

positive; also it is *not* finitely integrable at ∞. Therefore, as x travels from 0 to ∞, we see that $f(x)$ increases from 0 to ∞, thus f maps $[0, \infty)$ to $[0, \infty)$. On the other hand, when $z = x$ is negative, then

$$f'(z) = \alpha|x|^{\alpha-1}e^{i\pi(\alpha-1)} = -\alpha|x|^{\alpha-1}e^{i\pi\alpha},$$

so f maps the segment $(-\infty, 0]$ to $(e^{i\pi\alpha}\infty, 0]$. The situation is illustrated in Figure 3 where the infinite segment A is mapped to A' and the segment B is mapped to B', with the direction of travel indicated in Figure 3.

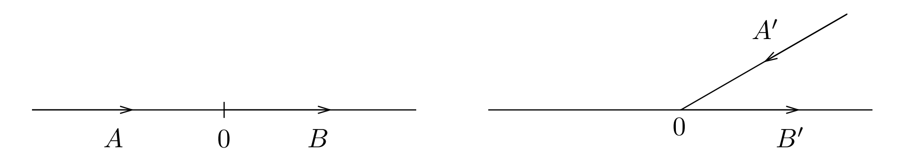


Figure 3. The conformal map z^α

Example 2. Next, we consider for $z \in \mathbb{H}$,

$$f(z) = \int_0^z \frac{d\zeta}{(1-\zeta^2)^{1/2}},$$

where the integral is taken from 0 to z along any path in the closed upper half-plane. We choose the branch for $(1-\zeta^2)^{1/2}$ that makes it holomorphic in the upper half-plane and positive when $-1 < \zeta < 1$. As a result

$$(1-\zeta^2)^{-1/2} = i(\zeta^2 - 1)^{-1/2} \quad \text{when } \zeta > 1.$$

We observe that f maps the real line to the boundary of the half-strip pictured in Figure 4.

In fact, since $f(\pm 1) = \pm\pi/2$, and $f'(x) > 0$ if $-1 < x < 1$, we see that f maps the segment B to B'. Moreover

$$f(x) = \frac{\pi}{2} + \int_1^x f'(x)\,dx \quad \text{when } x > 1, \quad \text{and} \quad \int_1^\infty \frac{dx}{(x^2-1)^{1/2}} = \infty.$$

Thus, as x travels along the segment C, the image traverses the infinite segment C'. Similarly segment A is mapped to A'.

Note the connection of this example with Example 8 in Section 1.2. In fact, one can show that the function $f(z)$ is the inverse to the function

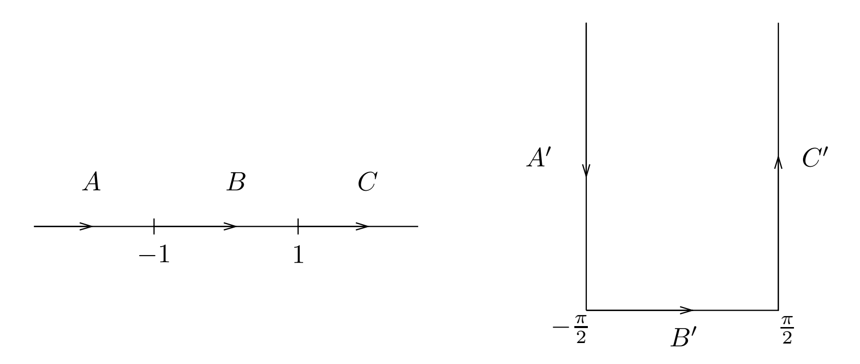


Figure 4. Mapping of the boundary in Example 2

$\sin z$, and hence f takes $\mathbb{H}$ conformally to the interior of the half-strip bounded by the segments A', B', and C'.

EXAMPLE 3. Here we take

$$f(z) = \int_0^z \frac{d\zeta}{[(1-\zeta^2)(1-k^2\zeta^2)]^{1/2}}, \qquad z \in \mathbb{H},$$

where k is a fixed real number with $0 < k < 1$ (the branch of $[(1-\zeta^2)\ (1-k^2\zeta^2)]^{1/2}$ in the upper half-plane is chosen to be the one that is positive when ζ is real and $-1 < \zeta < 1$). Integrals of this kind are called **elliptic integrals**, because variants of these arise in the calculation of the arc-length of an ellipse. We shall observe that f maps the real axis onto the rectangle shown in Figure 5(b), where K and K' are determined by

$$K = \int_0^1 \frac{dx}{[(1-x^2)(1-k^2x^2)]^{1/2}},\ K' = \int_1^{1/k} \frac{dx}{[(x^2-1)(1-k^2x^2)]^{1/2}}.$$

We divide the real axis into four "segments," with division points $-1/k$, -1, 1, and $1/k$ (see Figure 5(a)). The segments are $[-1/k, -1]$, $[-1, 1]$, $[1, 1/k]$, and $[1/k, -1/k]$, the last consisting of the union of the two half-segments $[1/k, \infty)$ and $(-\infty, -1/k]$. It is clear from the definitions that $f(\pm 1) = \pm K$, and since $f'(x) > 0$, when $-1 < x < 1$, it follows that f maps the segment $[-1, 1]$ to $[-K, K]$. Moreover, since

$$f(z) = K + \int_1^x \frac{d\zeta}{[(1-\zeta^2)(1-k^2\zeta^2)]^{1/2}} \qquad \text{if } 1 < x < 1/k,$$

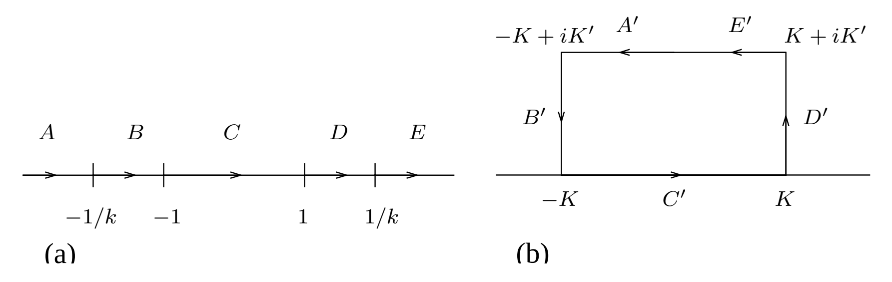


Figure 5. Mapping of the boundary in Example 3

we see that f maps the segment $[1, 1/k]$ to $[K, K + iK']$, where K' was defined above. Similarly, f maps $[-1/k, -1]$ to $[-K + iK', -K]$. Next, when $x > 1/k$ we have

$$f'(x) = -\frac{1}{[(x^2-1)(k^2x^2-1)]^{1/2}},$$

and therefore,

$$f(x) = K + iK' - \int_{1/k}^{x} \frac{dx}{[(x^2-1)(k^2x^2-1)]^{1/2}}.$$

However,

$$\int_{1/k}^{\infty} \frac{dx}{[(x^2-1)(k^2x^2-1)]^{1/2}} = \int_0^1 \frac{dx}{[(1-x^2)(1-k^2x^2)]^{1/2}},$$

as can be seen by making the change of variables $x = 1/ku$ in the integral on the left. Thus f maps the segment $[1/k, \infty)$ to the segment $[K + iK', iK')$. Similarly f maps $(-\infty, -1/k]$ to $[-K + iK', iK')$. Altogether, then, f maps the real axis to the above rectangle, with the point at infinity corresponding to the mid-point of the upper side of the rectangle.

The results obtained so far lead naturally to two problems.

The first, which we pursue next, consists of a generalization of the above examples. More precisely we define the Schwarz-Christoffel integral and prove that it maps the real line to a polygonal line.

Second, we note that in the examples above little was inferred about the behavior of f in $\mathbb{H}$ itself. In particular, we have not shown that f maps $\mathbb{H}$ conformally to the interior of the corresponding polygon. After a careful study of the boundary behavior of conformal maps, we

prove a theorem that guarantees that the conformal map from the upper half-plane to a simply connected region bounded by a polygonal line is essentially given by a Schwarz-Christoffel integral.

4.2 The Schwarz-Christoffel integral

With the examples of the previous section in mind, we define the general **Schwarz-Christoffel integral** by

$$S(z) = \int_0^z \frac{d\zeta}{(\zeta - A_1)^{\beta_1} \cdots (\zeta - A_n)^{\beta_n}}. \tag{5}$$

Here $A_1 < A_2 < \cdots < A_n$ are n distinct points on the real axis arranged in increasing order. The exponents β_k will be assumed to satisfy the conditions $\beta_k < 1$ for each k and $1 < \sum_{k=1}^n \beta_k$.[6]

The integrand in (5) is defined as follows: $(z - A_k)^{\beta_k}$ is that branch (defined in the complex plane slit along the infinite ray $\{A_k + iy : y \leq 0\}$) which is positive when $z = x$ is real and $x > A_k$. As a result

$$(z - A_k)^{\beta_k} = \begin{cases} (x - A_k)^{\beta_k} & \text{if } x \text{ is real and } x > A_k, \\ |x - A_k|^{\beta_k} e^{i\pi\beta_k} & \text{if } x \text{ is real and } x < A_k. \end{cases}$$

The complex plane slit along the union of the rays $\cup_{k=1}^n \{A_k + iy : y \leq 0\}$ is simply connected (see Exercise 19), so the integral that defines $S(z)$ is holomorphic in this open set. Since the requirement $\beta_k < 1$ implies that the singularities $(\zeta - A_k)^{-\beta_k}$ are integrable near A_k, the function S is continuous up to the real line, including the points A_k, with $k = 1, \ldots, n$. Finally, this continuity condition implies that the integral can be taken along any path in the complex plane that avoids the union of the *open* slits $\cup_{k=1}^n \{A_k + iy : y < 0\}$.

Now

$$\left| \prod_{k=1}^n (\zeta - A_k)^{-\beta_k} \right| \leq c|\zeta|^{-\sum \beta_k}$$

for $|\zeta|$ large, so the assumption $\sum \beta_k > 1$ guarantees the convergence of the integral (5) at infinity. This fact and Cauchy's theorem imply that $\lim_{r\to\infty} S(re^{i\theta})$ exists and is independent of the angle θ, $0 \leq \theta \leq \pi$. We call this limit a_∞, and we let $a_k = S(A_k)$ for $k = 1, \ldots, n$.

[6]Note that the case $\sum \beta_k \leq 1$, which occurs in Examples 1 and 2 above is excluded. However, a modification of the proposition that follows can be made to take these cases into account; but then $S(z)$ is no longer bounded in the upper half-plane.

Proposition 4.1 *Suppose $S(z)$ is given by* (5).

(i) *If $\sum_{k=1}^{n} \beta_k = 2$, and $\mathfrak{p}$ denotes the polygon whose vertices are given (in order) by $a_1, \ldots, a_n$, then S maps the real axis onto $\mathfrak{p} - \{a_\infty\}$. The point a_∞ lies on the segment $[a_n, a_1]$ and is the image of the point at infinity. Moreover, the (interior) angle at the vertex a_k is $\alpha_k \pi$ where $\alpha_k = 1 - \beta_k$.*

(ii) *There is a similar conclusion when $1 < \sum_{k=1}^{n} \beta_k < 2$, except now the image of the extended line is the polygon of $n+1$ sides with vertices $a_1, a_2, \ldots, a_n, a_\infty$. The angle at the vertex a_∞ is $\alpha_\infty \pi$ with $\alpha_\infty = 1 - \beta_\infty$, where $\beta_\infty = 2 - \sum_{k=1}^{n} \beta_k$.*

Figure 6 illustrates the proposition. The idea of the proof is already captured in Example 1 above.

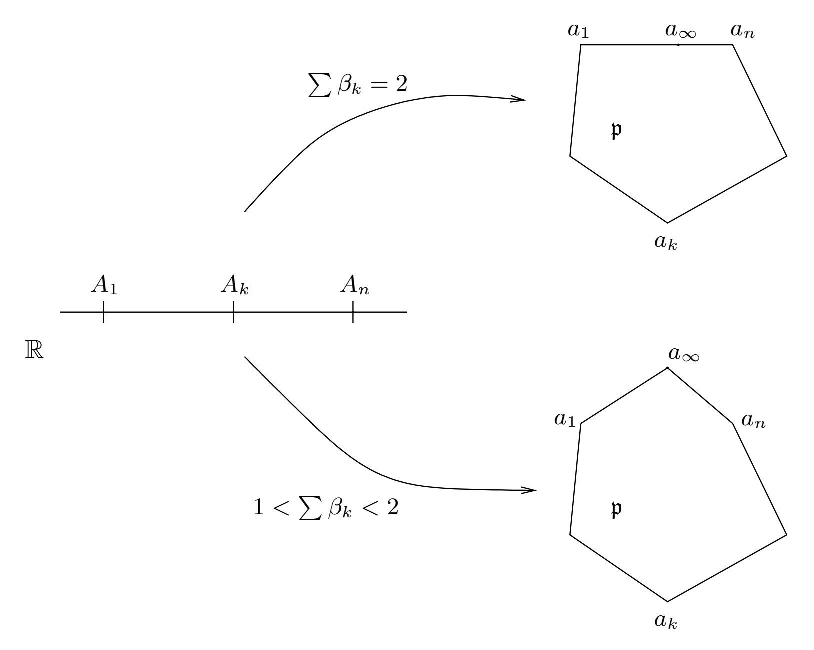


Figure 6. Action of the integral $S(z)$

Proof. We assume that $\sum_{k=1}^{n} \beta_k = 2$. If $A_k < x < A_{k+1}$ when $1 \leq k \leq n-1$, then

$$S'(x) = \prod_{j \leq k} (x - A_j)^{-\beta_j} \prod_{j > k} (x - A_j)^{-\beta_j}.$$

Hence

$$\arg S'(x) = \arg\left(\prod_{j>k}(x-A_j)^{-\beta_j}\right) = \arg\prod_{j>k} e^{-i\pi\beta_j} = -\pi\sum_{j>k}\beta_j,$$

which of course is constant when x traverses the interval (A_k, A_{k+1}). Since

$$S(x) = S(A_k) + \int_{A_k}^{x} S'(y)\,dy,$$

we see that as x varies from A_k to A_{k+1}, $S(x)$ varies from $S(A_k) = a_k$ to $S(A_{k+1}) = a_{k+1}$ along the straight line segment[7] $[a_k, a_{k+1}]$, and this makes an angle of $-\pi\sum_{j>k}\beta_j$ with the real axis. Similarly, when $A_n < x$ then $S'(x)$ is positive, while if $x < A_1$, the argument of $S'(x)$ is $-\pi\sum_{k=1}^{n}\beta_k = -2\pi$, and so $S'(x)$ is again positive. Thus as x varies on $[A_n, +\infty)$, $S(x)$ varies along a straight line (parallel to the x-axis) between a_n and a_∞; similarly $S(x)$ varies along a straight line (parallel to that axis) between a_∞ and a_1 as x varies in $(-\infty, A_1]$. Moreover, the union of $[a_n, a_\infty)$ and $(a_\infty, a_1]$ is the segment $[a_n, a_1]$ with the point a_∞ removed.

Now the increase of the angle of $[a_{k+1}, a_k]$ over that of $[a_{k-1}, a_k]$ is $\pi\beta_k$, which means that the angle at the vertex a_k is $\pi\alpha_k$. The proof when $1 < \sum_{k=1}^{n}\beta_k < 2$ is similar, and is left to the reader.

As elegant as this proposition is, it does not settle the problem of finding a conformal map from the half-plane to a given region P that is bounded by a polygon. There are two reasons for this.

1. It is not true for general n and generic choices of $A_1, \ldots, A_n$, that the polygon (which is the image of the real axis under S) is simple, that is, it does not cross itself. Nor is it true in general that the mapping S is conformal on the upper half-plane.

2. Neither does the proposition show that starting with a simply connected region P (whose boundary is a polygonal line $\mathfrak{p}$) the mapping S is, for certain choices of $A_1, \ldots, A_n$ and simple modifications, a conformal map from $\mathbb{H}$ to P. That however is the case, and is the result whose proof we now turn to.

[7] We denote the closed straight line segment between two complex numbers z and w by $[z, w]$, that is, $[z, w] = \{(1-t)z + tw : t \in [0,1]\}$. If we restrict $0 < t < 1$, then (z, w) denotes the open line segment between z and w. Similarly for the half-open segments $[z, w)$ and $(z, w]$ obtained by restricting $0 \le t < 1$ and $0 < t \le 1$, respectively.

4.3 Boundary behavior

In what follows we shall consider a **polygonal region** P, namely a bounded, simply connected open set whose boundary is a polygonal line $\mathfrak{p}$. In this context, we always assume that the polygonal line is closed, and we sometimes refer to $\mathfrak{p}$ as a **polygon**.

To study conformal maps from the half-plane $\mathbb{H}$ to P, we consider first the conformal maps from the disc $\mathbb{D}$ to P, and their boundary behavior.

Theorem 4.2 *If $F : \mathbb{D} \to P$ is a conformal map, then F extends to a continuous bijection from the closure $\overline{\mathbb{D}}$ of the disc to the closure $\overline{P}$ of the polygonal region. In particular, F gives rise to a bijection from the boundary of the disc to the boundary polygon $\mathfrak{p}$.*

The main point consists in showing that if z_0 belongs to the unit circle, then $\lim_{z \to z_0} F(z)$ exists. To prove this, we need a preliminary result, which uses the fact that if $f : U \to f(U)$ is conformal, then

$$\text{Area}(f(U)) = \iint_U |f'(z)|^2 \, dx \, dy.$$

This assertion follows from the definition, $\text{Area}(f(U)) = \iint_{f(U)} dx \, dy$, and the fact that the determinant of the Jacobian in the change of variables $w = f(z)$ is simply $|f'(z)|^2$, an observation we made in equation (4), Section 2.2, Chapter 1.

Lemma 4.3 *For each $0 < r < 1/2$, let C_r denote the circle centered at z_0 of radius r. Suppose that for all sufficiently small r we are given two points z_r and z_r' in the unit disc that also lie on C_r. If we let $\rho(r) = |f(z_r) - f(z_r')|$, then there exists a sequence $\{r_n\}$ of radii that tends to zero, and such that $\lim_{n\to\infty} \rho(r_n) = 0$.*

Proof. If not, there exist $0 < c$ and $0 < R < 1/2$ such that $c \leq \rho(r)$ for all $0 < r \leq R$. Observe that

$$f(z_r) - f(z_r') = \int_\alpha f'(\zeta) \, d\zeta,$$

where the integral is taken over the arc α on C_r that joins z_r and z_r' in $\mathbb{D}$. If we parametrize this arc by $z_0 + re^{i\theta}$ with $\theta_1(r) \leq \theta \leq \theta_2(r)$, then

$$\rho(r) \leq \int_{\theta_1(r)}^{\theta_2(r)} |f'(z)| r \, d\theta.$$

We now apply the Cauchy-Schwarz inequality to see that

$$\rho(r) \leq \left(\int_{\theta_1(r)}^{\theta_2(r)} |f'(z)|^2 r \, d\theta \right)^{1/2} \left(\int_{\theta_1(r)}^{\theta_2(r)} r \, d\theta \right)^{1/2}.$$

Squaring both sides and dividing by r yields

$$\frac{\rho(r)^2}{r} \leq 2\pi \int_{\theta_1(r)}^{\theta_2(r)} |f'(z)|^2 r \, d\theta.$$

We may now integrate both sides from 0 to R, and since $c \leq \rho(r)$ on that region we obtain

$$c^2 \int_0^R \frac{dr}{r} \leq 2\pi \int_0^R \int_{\theta_1(r)}^{\theta_2(r)} |f'(z)|^2 r \, d\theta dr \leq 2\pi \int\!\!\int_{\mathbb{D}} |f'(z)|^2 \, dxdy.$$

Now the left-hand side is infinite because $1/r$ is not integrable near the origin, and the right-hand side is bounded because the area of the polygonal region is bounded, so this yields the desired contradiction and concludes the proof of the lemma.

Lemma 4.4 *Let z_0 be a point on the unit circle. Then $F(z)$ tends to a limit as z approaches z_0 within the unit disc.*

Proof. If not, there are two sequences $\{z_1, z_2, \ldots\}$ and $\{z_1', z_2', \ldots\}$ in the unit disc that converge to z_0 and are so that $F(z_n)$ and $F(z_n')$ converge to two distinct points ζ and ζ' in the closure of P. Since F is conformal, the points ζ and ζ' must lie on the boundary $\mathfrak{p}$ of P. We may therefore choose two disjoint discs D and D' centered at ζ and ζ', respectively, that are at a distance $d > 0$ from each other. For all large n, $F(z_n) \in D$ and $F(z_n') \in D'$. Therefore, there exist two continuous curves[8] Λ and Λ' in $D \cap P$ and $D' \cap P$, respectively, with $F(z_n) \in \Lambda$ and $F(z_n') \in \Lambda'$ for all large n, and with the end-points of Λ and Λ' equal to ζ and ζ', respectively.

Define $\lambda = F^{-1}(\Lambda)$ and $\lambda' = F^{-1}(\Lambda')$. Then λ and λ' are two continuous curves in $\mathbb{D}$. Moreover, both λ and λ' contain infinitely many points in each sequence $\{z_n\}$ and $\{z_n'\}$. Recall that these sequences converge to z_0. By continuity, the circle C_r centered at z_0 and of radius r will intersect λ and λ' for all small r, say at some points $z_r \in \lambda$ and $z_r' \in \lambda'$. This

[8]By a continuous curve, we mean the image of a continuous (not necessarily piecewise-smooth) function from a closed interval $[a, b]$ to $\mathbb{C}$.

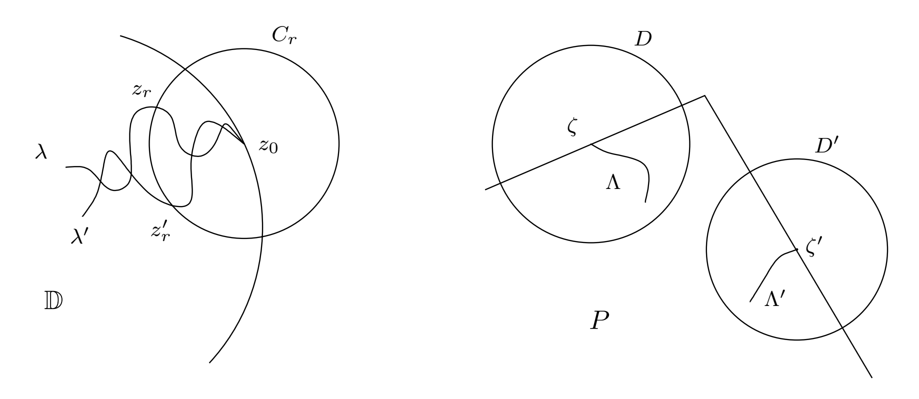


Figure 7. Illustration for the proof of Lemma 4.4

contradicts the previous lemma, because $|F(z_r) - F(z'_r)| > d$. Therefore $F(z)$ converges to a limit on $\mathfrak{p}$ as z approaches z_0 from within the unit disc, and the proof is complete.

Lemma 4.5 *The conformal map F extends to a continuous function from the closure of the disc to the closure of the polygon.*

Proof. By the previous lemma, the limit

$$\lim_{z \to z_0} F(z)$$

exists, and we define $F(z_0)$ to be the value of this limit. There remains to prove that F is continuous on the closure of the unit disc. Given ϵ, there exists δ such that whenever $z \in \mathbb{D}$ and $|z - z_0| < \delta$, then $|F(z) - F(z_0)| < \epsilon$. Now if z belongs to the boundary of $\mathbb{D}$ and $|z - z_0| < \delta$, then we may choose w such that $|F(z) - F(w)| < \epsilon$ and $|w - z_0| < \delta$. Therefore

$$|F(z) - F(z_0)| \leq |F(z) - F(w)| + |F(w) - F(z_0)| < 2\epsilon,$$

and the lemma is established.

We may now complete the proof of the theorem. We have shown that F extends to a continuous function from $\overline{\mathbb{D}}$ to $\overline{P}$. The previous argument can be applied to the inverse G of F. Indeed, the key geometric property of the unit disc that we used was that if z_0 belongs to the boundary of $\mathbb{D}$, and C is any small circle centered at z_0, then $C \cap \mathbb{D}$ consists of an arc. Clearly, this property also holds at every boundary point of the

polygonal region P. Therefore, G also extends to a continuous function from $\overline{P}$ to $\overline{\mathbb{D}}$. It suffices to now prove that the extensions of F and G are inverses of each other. If $z \in \partial\mathbb{D}$ and $\{z_k\}$ is a sequence in the disc that converges to z, then $G(F(z_k)) = z_k$, so after taking the limit and using the fact that F is continuous, we conclude that $G(F(z)) = z$ for all $z \in \overline{\mathbb{D}}$. Similarly, $F(G(w)) = w$ for all $w \in \overline{P}$, and the theorem is proved.

The circle of ideas used in this proof can be used to prove more general theorems on the boundary continuity of conformal maps. See Exercise 18 and Problem 6 below.

4.4 The mapping formula

Suppose P is a polygonal region bounded by a polygon $\mathfrak{p}$ whose vertices are ordered consecutively $a_1, a_2, \ldots, a_n$, and with $n \geq 3$. We denote by $\pi\alpha_k$ the interior angle at a_k, and define the exterior angle $\pi\beta_k$ by $\alpha_k + \beta_k = 1$. A simple geometric argument provides $\sum_{k=1}^n \beta_k = 2$.

We shall consider conformal mappings of the half-plane $\mathbb{H}$ to P, and make use of the results of the previous section regarding conformal maps from the disc $\mathbb{D}$ to P. The standard correspondences $w = (i - z)/(i + z)$, $z = i(1 - w)/(1 + w)$ allows us to go back and forth between $z \in \mathbb{H}$ and $w \in \mathbb{D}$. Notice that the boundary point $w = -1$ of the circle corresponds to the point at infinity on the line, and so the conformal map of $\mathbb{H}$ to $\mathbb{D}$ extends to a continuous bijection of the boundary of $\mathbb{H}$, which for the purpose of this discussion includes the point at infinity.

Let F be a conformal map from $\mathbb{H}$ to P. (Its existence is guaranteed by the Riemann mapping theorem and the previous discussion.) We assume first that none of the vertices of $\mathfrak{p}$ correspond to the point at infinity. Therefore, there are real numbers $A_1, A_2, \ldots, A_n$ so that $F(A_k) = a_k$ for all k. Since F is continuous and injective, and the vertices are numbered consecutively, we may conclude that the A_k's are in either increasing or decreasing order. After relabeling the vertices a_k and the points A_k, we may assume that $A_1 < A_2 < \cdots < A_n$. These points divide the real line into $n - 1$ segments $[A_k, A_{k+1}]$, $1 \leq k \leq n - 1$, and the segment that consists of the join of the two half-segments $(-\infty, A_1] \cup [A_n, \infty)$. These are mapped bijectively onto the corresponding sides of the polygon, that is, the segments $[a_k, a_{k+1}]$, $1 \leq k \leq n - 1$, and $[a_n, a_1]$ (see Figure 8).

Theorem 4.6 *There exist complex numbers c_1 and c_2 so that the conformal map F of $\mathbb{H}$ to P is given by*

$$F(z) = c_1 S(z) + c_2$$

where S is the Schwarz-Christoffel integral introduced in Section 4.2.

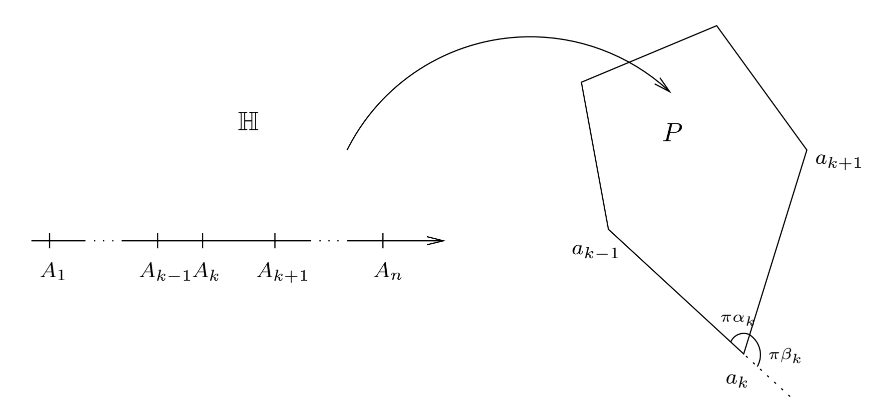


Figure 8. The mapping F

Proof. We first consider z in the upper half-plane lying above the two adjacent segments $[A_{k-1}, A_k]$ and $[A_k, A_{k+1}]$, where $1 < k < n$. We note that F maps these two segments to two segments that intersect at $a_k = F(A_k)$ at an angle $\pi\alpha_k$.

By choosing a branch of the logarithm we can in turn define

$$h_k(z) = (F(z) - a_k)^{1/\alpha_k}$$

for all z in the half-strip in the upper half-plane bounded by the lines $\mathrm{Re}(z) = A_{k-1}$ and $\mathrm{Re}(z) = A_{k+1}$. Since F continues to the boundary of $\mathbb{H}$, the map h_k is actually continuous up to the segment (A_{k-1}, A_{k+1}) on the real line. By construction h_k will map the segment $[A_{k-1}, A_{k+1}]$ to a (straight) segment L_k in the complex plane, with A_k mapped to 0. We may therefore apply the Schwarz reflection principle to see that h_k is analytically continuable to a holomorphic function in the two-way infinite strip $A_{k-1} < \mathrm{Re}(z) < A_{k+1}$ (see Figure 9). We claim that h_k' never vanishes in that strip. First, if z belongs to the open upper half-strip, then

$$\frac{F'(z)}{F(z) - F(A_k)} = \alpha_k \, \frac{h_k'(z)}{h_k(z)},$$

and since F is conformal, we have $F'(z) \neq 0$ so $h_k'(z) \neq 0$ (Proposition 1.1). By reflection, this also holds in the lower half-strip, and it remains to investigate points on the segment (A_{k-1}, A_{k+1}). If $A_{k-1} < x < A_{k+1}$, we note that the image under h_k of a small half-disc centered at x and contained in $\mathbb{H}$ lies on one side of the straight line segment

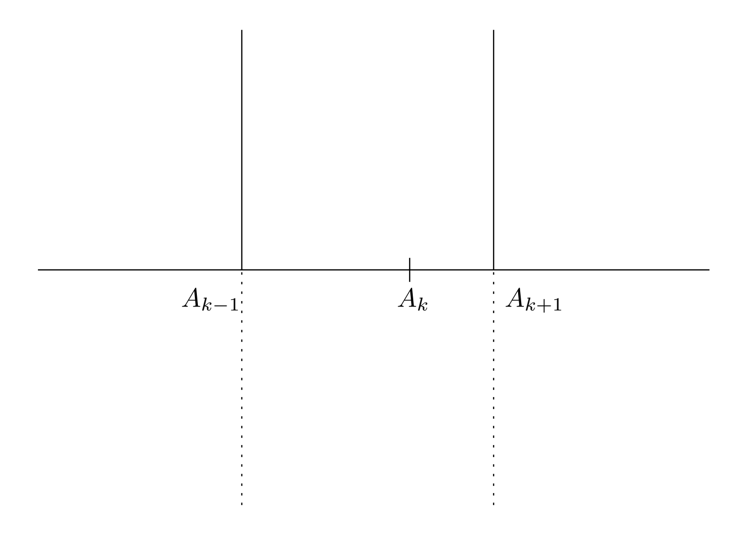


Figure 9. Schwarz reflection

L_k. Since h_k is injective up to L_k (because F is) the symmetry in the Schwarz reflection principle guarantees that h_k is injective in the whole disc centered at x, whence $h_k'(x) \neq 0$, whence $h_k'(z) \neq 0$ for all z in the strip $A_{k-1} < \operatorname{Re}(z) < A_{k+1}$.

Now because $F' = \alpha_k h_k^{-\beta_k} h_k'$ and $F'' = -\beta_k \alpha_k h_k^{-\beta_k - 1}(h_k')^2 + \alpha_k h_k^{-\beta_k} h_k''$, the fact that $h_k'(z) \neq 0$ implies that

$$\frac{F''(z)}{F'(z)} = \frac{-\beta_k}{z - A_k} + E_k(z),$$

where E_k is holomorphic in the strip $A_{k-1} < \operatorname{Re}(z) < A_{k+1}$. A similar result holds for $k = 1$ and $k = n$, namely

$$\frac{F''(z)}{F'(z)} = -\frac{\beta_1}{z - A_1} + E_1,$$

where E_1 is holomorphic in the strip $-\infty < \operatorname{Re}(z) < A_2$, and

$$\frac{F''(z)}{F'(z)} = -\frac{\beta_n}{z - A_n} + E_n,$$

where E_n is holomorphic in the strip $A_{n-1} < \operatorname{Re}(z) < \infty$. Finally, another application of the reflection principle shows that F is continuable in the exterior of a disc $|z| \leq R$, for large R (say $R > \max_{1 \leq k \leq n} |A_k|$). Indeed, we may continue F across the union of the segments

$(-\infty, A_1) \cup (A_n, \infty)$ since their image under F is a straight line segment and Schwarz reflection applies. The fact that F maps the upper half-plane to a bounded region shows that the analytic continuation of F outside a large disc is also bounded, and hence holomorphic at infinity. Thus F''/F' is holomorphic at infinity and we claim that it goes to 0 as $|z| \to \infty$. Indeed, we may expand F at $z = \infty$ as

$$F(z) = c_0 + \frac{c_1}{z} + \frac{c_2}{z^2} + \cdots .$$

This after differentiation shows that F''/F' decays like $1/z$ as $|z|$ becomes large, and proves our claim.

Altogether then, because the various strips overlap and cover the entire complex plane,

$$\frac{F''(z)}{F'(z)} + \sum_{k=1}^{n} \frac{\beta_k}{z - A_k}$$

is holomorphic in the entire plane and vanishes at infinity; thus, by Liouville's theorem it is zero. Hence

$$\frac{F''(z)}{F'(z)} = -\sum_{k=1}^{n} \frac{\beta_k}{z - A_k}.$$

From this we contend that $F'(z) = c(z - A_1)^{-\beta_1} \cdots (z - A_n)^{-\beta_n}$. Indeed, denoting this product by $Q(z)$, we have

$$\frac{Q'(z)}{Q(z)} = -\sum_{k=1}^{n} \frac{\beta_k}{z - A_k}.$$

Therefore

$$\frac{d}{dz}\left(\frac{F'(z)}{Q(z)}\right) = 0,$$

which proves the contention. A final integration yields the theorem.

We may now withdraw the hypothesis we made at the beginning that F did not map the point at infinity to a vertex of P, and obtain a formula for that case as well.

Theorem 4.7 *If F is a conformal map from the upper half-plane to the polygonal region P and maps the points $A_1, \ldots, A_{n-1}, \infty$ to the vertices of $\mathfrak{p}$, then there exist constants C_1 and C_2 such that*

$$F(z) = C_1 \int_0^z \frac{d\zeta}{(\zeta - A_1)^{\beta_1} \cdots (\zeta - A_{n-1})^{\beta_{n-1}}} + C_2.$$

In other words, the formula is obtained by deleting the last term in the Schwarz-Christoffel integral (5).

Proof. After a preliminary translation, we may assume that $A_j \neq 0$ for $j = 1, \ldots, n-1$. Choose a point $A_n^* > 0$ on the real line, and consider the fractional linear map defined by

$$\Phi(z) = A_n^* - \frac{1}{z}.$$

Then Φ is an automorphism of the upper half-plane. Let $A_k^* = \Phi(A_k)$ for $k = 1, \ldots, n-1$, and note that $A_n^* = \Phi(\infty)$. Then

$$(F \circ \Phi^{-1})(A_k^*) = a_k \quad \text{for all } k = 1, 2, \ldots, n.$$

We can now apply the Schwarz-Christoffel formula just proved to find that

$$(F \circ \Phi^{-1})(z') = C_1 \int_0^{z'} \frac{d\zeta}{(\zeta - A_1^*)^{\beta_1} \cdots (\zeta - A_n^*)^{\beta_n}} + C_2.$$

The change of variables $\zeta = \Phi(w)$ satisfies $d\zeta = dw/w^2$, and since we can write $2 = \beta_1 + \cdots + \beta_n$, we obtain

$$\begin{aligned}(F \circ \Phi^{-1})(z') &= C_1 \int_0^{\Phi^{-1}(z')} \frac{dw}{(w(A_n^* - A_1^*) - 1)^{\beta_1} \cdots (w(A_n^* - A_{n-1}^*) - 1)^{\beta_{n-1}}} + C_2' \\ &= C_1' \int_0^{\Phi^{-1}(z')} \frac{dw}{(w - 1/(A_n^* - A_1^*))^{\beta_1} \cdots (w - 1/(A_n^* - A_{n-1}^*))^{\beta_{n-1}}} + C_2'.\end{aligned}$$

Finally, we note that $1/(A_n^* - A_k^*) = A_k$ and set $\Phi^{-1}(z') = z$ in the above equation to conclude that

$$F(z) = C_1' \int_0^z \frac{dw}{(w - A_1)^{\beta_1} \cdots (w - A_{n-1})^{\beta_{n-1}}} + C_2',$$

as was to be shown.

4.5 Return to elliptic integrals

We consider again the elliptic integral

$$I(z) = \int_0^z \frac{d\zeta}{[(1 - \zeta^2)(1 - k^2\zeta^2)]^{1/2}} \quad \text{with } 0 < k < 1,$$

which arose in Example 3 of Section 4.1. We saw that it mapped the real axis to the rectangle R with vertices $-K$, K, $K+iK'$, and $-K+iK'$. We will now see that this mapping is a conformal mapping of $\mathbb{H}$ to the interior of R.

According to Theorem 4.6 there is a conformal map F to the rectangle, that maps four points on the real axis to the vertices of R. By preceding this map with a suitable automorphism of $\mathbb{H}$ we may assume that F maps -1, 0, 1 to $-K$, 0, K, respectively. Indeed, by using a preliminary automorphism, we may assume that $-K$, 0, K are the images of points A_1, 0, A_2 with $A_1 < 0 < A_2$; then we can further take $A_1 = -1$ and $A_2 = 1$. See Exercise 15.

Next, let ℓ be chosen with $0 < \ell < 1$, so that $1/\ell$ is the point on the real line mapped by F to the vertex $K+iK'$, which is the vertex next in order after $-K$ and K. We claim that $F(-1/\ell)$ is the vertex $-K+iK'$. Indeed, if $F^*(z) = -F(-\overline{z})$, then by the symmetry of R, F^* is also a conformal map of $\mathbb{H}$ to R; moreover $F^*(0) = 0$, and $F^*(\pm 1) = \pm K$. Thus $F^{-1} \circ F^*$ is an automorphism of $\mathbb{H}$ that fixes the points -1, 0, and 1. Hence $F^{-1} \circ F^*$ is the identity (see Exercise 15), and $F = F^*$, from which it follows that

$$F(-1/\ell) = -\overline{F}(1/\ell) = -K + iK'.$$

Therefore, by Theorem 4.6

$$F(z) = c_1 \int_0^z \frac{d\zeta}{[(1-\zeta^2)(1-\ell^2\zeta^2)]^{1/2}} + c_2.$$

Setting $z = 0$ gives $c_2 = 0$, and letting $z = 1$, $z = 1/\ell$, yields

$$K(k) = c_1 K(\ell) \quad \text{and} \quad K'(k) = c_1 K'(\ell),$$

where

$$K(k) = \int_0^1 \frac{dx}{[(1-x^2)(1-k^2x^2)]^{1/2}},$$

$$K'(k) = \int_1^{1/k} \frac{dx}{[(x^2-1)(1-k^2x^2)]^{1/2}}.$$

Now $K(k)$ is clearly strictly increasing as k varies in $(0,1)$. Moreover, a change of variables (Exercise 24) establishes the identity

$$K'(k) = K(\tilde{k}) \quad \text{where } \tilde{k}^2 = 1 - k^2 \text{ and } \tilde{k} > 0,$$

and this shows that $K'(k)$ is strictly decreasing. Hence $K(k)/K'(k)$ is strictly increasing. Since $K(k)/K'(k) = K(\ell)/K'(\ell)$, we must have $k = \ell$, and finally $c_1 = 1$. This shows that $I(z) = F(z)$, and hence I is conformal, as was to be proved.

A final observation is of significance. A basic insight into elliptic integrals is obtained by passing to their inverse functions. We therefore consider $z \mapsto \mathrm{sn}(z)$, the inverse map of $z \mapsto I(z)$.[9] It transforms the closed rectangle into the closed upper half-plane. Now consider the series of rectangles $R = R_0,\ R_1,\ R_2, \ldots$ gotten by reflecting successively along the lower sides (Figure 10).

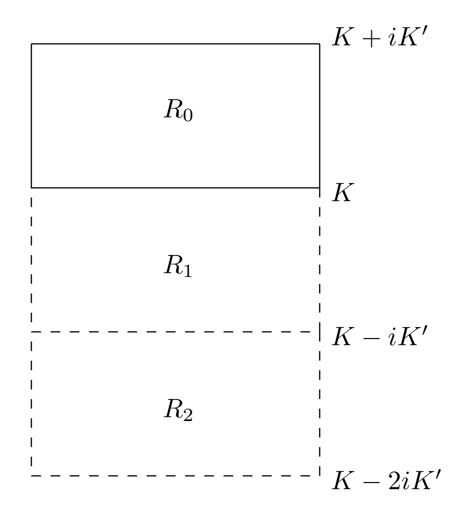


Figure 10. Reflections of $R = R_0$

With $\mathrm{sn}(z)$ defined in R_0, we can by the reflection principle extend it to R_1 by setting $\mathrm{sn}(z) = \overline{\mathrm{sn}(\overline{z})}$ whenever $z \in R_1$ (note that then $\overline{z} \in R_0$). Next we can extend $\mathrm{sn}(z)$ to R_2 by setting $\mathrm{sn}(z) = \overline{\mathrm{sn}(-iK' + \overline{z})}$ if $z \in R_2$ and noting that if $z \in R_2$, then $-iK' + \overline{z} \in R_1$. Combining these reflections and continuing this way we see that we can extend $\mathrm{sn}(z)$ in the entire strip $-K < \mathrm{Re}(z) < K$, so that $\mathrm{sn}(z) = \mathrm{sn}(z + 2iK')$.

Similarly, by reflecting in a series of horizontal rectangles, and combining these with the previous reflections, we see that $\mathrm{sn}(z)$ can be continued to the complex plane and also satisfies $\mathrm{sn}(z) = \mathrm{sn}(z + 4K)$. Thus $\mathrm{sn}(z)$ is *doubly periodic* (with periods $4K$ and $2iK'$). A further examination

[9]The notation $\mathrm{sn}(z)$ in somewhat different form is due to Jacobi, and was adopted because of the analogy with $\sin z$.

shows that the only singularities $\text{sn}(z)$ are poles. Functions of this type, called "elliptic functions," are the subject of the next chapter.

5 Exercises

1. A holomorphic mapping $f : U \to V$ is a **local bijection** on U if for every $z \in U$ there exists an open disc $D \subset U$ centered at z, so that $f : D \to f(D)$ is a bijection.

Prove that a holomorphic map $f : U \to V$ is a local bijection on U if and only if $f'(z) \neq 0$ for all $z \in U$.

[Hint: Use Rouché's theorem as in the proof of Proposition 1.1.]

2. Supppose $F(z)$ is holomorphic near $z = z_0$ and $F(z_0) = F'(z_0) = 0$, while $F''(z_0) \neq 0$. Show that there are two curves Γ_1 and Γ_2 that pass through z_0, are orthogonal at z_0, and so that F restricted to Γ_1 is real and has a minimum at z_0, while F restricted to Γ_2 is also real but has a maximum at z_0.

[Hint: Write $F(z) = (g(z))^2$ for z near z_0, and consider the mapping $z \mapsto g(z)$ and its inverse.]

3. Suppose U and V are conformally equivalent. Prove that if U is simply connected, then so is V. Note that this conclusion remains valid if we merely assume that there exists a continuous bijection between U and V.

4. Does there exist a holomorphic surjection from the unit disc to $\mathbb{C}$?

[Hint: Move the upper half-plane "down" and then square it to get $\mathbb{C}$.]

5. Prove that $f(z) = -\frac{1}{2}(z + 1/z)$ is a conformal map from the half-disc $\{z = x + iy : |z| < 1,\ y > 0\}$ to the upper half-plane.

[Hint: The equation $f(z) = w$ reduces to the quadratic equation $z^2 + 2wz + 1 = 0$, which has two distinct roots in $\mathbb{C}$ whenever $w \neq \pm 1$. This is certainly the case if $w \in \mathbb{H}$.]

6. Give another proof of Lemma 1.3 by showing directly that the Laplacian of $u \circ F$ is zero.

[Hint: The real and imaginary parts of F satisfy the Cauchy-Riemann equations.]

7. Provide all the details in the proof of the formula for the solution of the Dirichlet problem in a strip discussed in Section 1.3. Recall that it suffices to compute the solution at the points $z = iy$ with $0 < y < 1$.

(a) Show that if $re^{i\theta} = G(iy)$, then

$$re^{i\theta} = i\frac{\cos \pi y}{1 + \sin \pi y}.$$

This leads to two separate cases: either $0 < y \leq 1/2$ and $\theta = \pi/2$, or $1/2 \leq$

$y < 1$ and $\theta = -\pi/2$. In either case, show that

$$r^2 = \frac{1 - \sin \pi y}{1 + \sin \pi y} \quad \text{and} \quad P_r(\theta - \varphi) = \frac{\sin \pi y}{1 - \cos \pi y \sin \varphi}.$$

(b) In the integral $\frac{1}{2\pi} \int_0^\pi P_r(\theta - \varphi)\tilde{f}_0(\varphi)\, d\varphi$ make the change of variables $t = F(e^{i\varphi})$. Observe that

$$e^{i\varphi} = \frac{i - e^{\pi t}}{i + e^{\pi t}},$$

and then take the imaginary part and differentiate both sides to establish the two identities

$$\sin \varphi = \frac{1}{\cosh \pi t} \quad \text{and} \quad \frac{d\varphi}{dt} = \frac{\pi}{\cosh \pi t}.$$

Hence deduce that

$$\begin{aligned} \frac{1}{2\pi} \int_0^\pi P_r(\theta - \varphi)\tilde{f}_0(\varphi)\, d\varphi &= \frac{1}{2\pi} \int_0^\pi \frac{\sin \pi y}{1 - \cos \pi y \sin \varphi} \tilde{f}_0(\varphi)\, d\varphi \\ &= \frac{\sin \pi y}{2} \int_{-\infty}^{\infty} \frac{f_0(t)}{\cosh \pi t - \cos \pi y}\, dt. \end{aligned}$$

(c) Use a similar argument to prove the formula for the integral $\frac{1}{2\pi} \int_{-\pi}^{0} P_r(\theta - \varphi)\tilde{f}_1(\varphi)\, d\varphi$.

8. Find a harmonic function u in the open first quadrant that extends continuously up to the boundary except at the points 0 and 1, and that takes on the following boundary values: $u(x, y) = 1$ on the half-lines $\{y = 0,\ x > 1\}$ and $\{x = 0,\ y > 0\}$, and $u(x, y) = 0$ on the segment $\{0 < x < 1, y = 0\}$.

[Hint: Find conformal maps $F_1, F_2, \ldots, F_5$ indicated in Figure 11. Note that $\frac{1}{\pi} \arg(z)$ is harmonic on the upper half-plane, equals 0 on the positive real axis, and 1 on the negative real axis.]

9. Prove that the function u defined by

$$u(x, y) = \text{Re}\left(\frac{i + z}{i - z}\right) \quad \text{and} \quad u(0, 1) = 0$$

is harmonic in the unit disc and vanishes on its boundary. Note that u is not bounded in $\mathbb{D}$.

10. Let $F : \mathbb{H} \to \mathbb{C}$ be a holomorphic function that satisfies

$$|F(z)| \leq 1 \quad \text{and} \quad F(i) = 0.$$

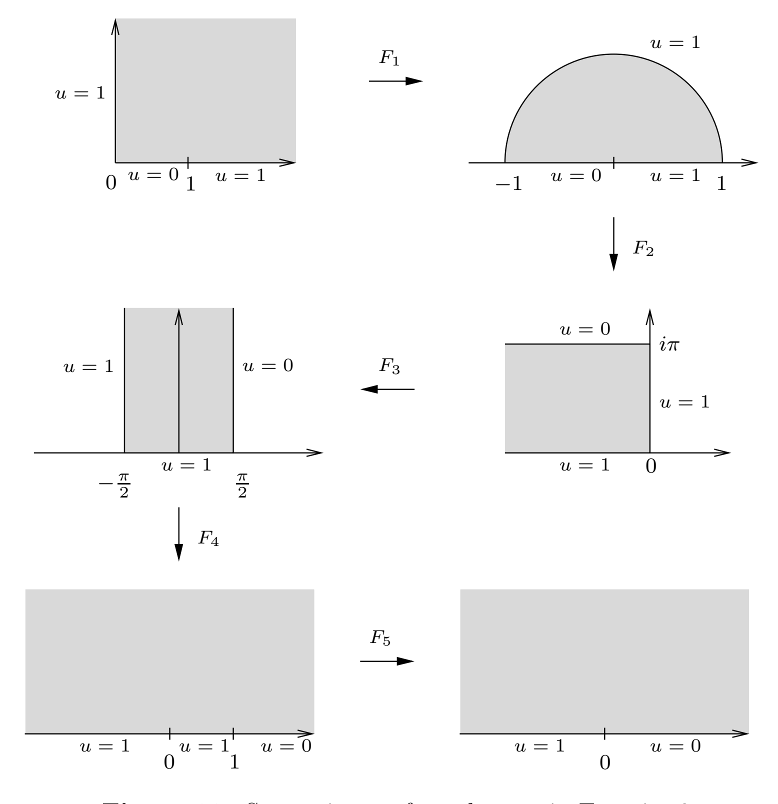


Figure 11. Successive conformal maps in Exercise 8

Prove that

$$|F(z)| \leq \left|\frac{z-i}{z+i}\right| \qquad \text{for all } z \in \mathbb{H}.$$

11. Show that if $f : D(0, R) \to \mathbb{C}$ is holomorphic, with $|f(z)| \leq M$ for some $M > 0$, then

$$\left|\frac{f(z) - f(0)}{M^2 - \overline{f(0)}f(z)}\right| \leq \frac{|z|}{MR}.$$

[Hint: Use the Schwarz lemma.]

12. A complex number $w \in \mathbb{D}$ is a **fixed point** for the map $f : \mathbb{D} \to \mathbb{D}$ if $f(w) = w$.

(a) Prove that if $f : \mathbb{D} \to \mathbb{D}$ is analytic and has two distinct fixed points, then f is the identity, that is, $f(z) = z$ for all $z \in \mathbb{D}$.

(b) Must every holomorphic function $f : \mathbb{D} \to \mathbb{D}$ have a fixed point? [Hint: Consider the upper half-plane.]

13. The **pseudo-hyperbolic distance** between two points $z, w \in \mathbb{D}$ is defined by

$$\rho(z,w) = \left| \frac{z-w}{1-\overline{w}z} \right|.$$

(a) Prove that if $f : \mathbb{D} \to \mathbb{D}$ is holomorphic, then

$$\rho(f(z), f(w)) \leq \rho(z,w) \qquad \text{for all } z, w \in \mathbb{D}.$$

Moreover, prove that if f is an automorphism of $\mathbb{D}$ then f preserves the pseudo-hyperbolic distance

$$\rho(f(z), f(w)) = \rho(z,w) \qquad \text{for all } z, w \in \mathbb{D}.$$

[Hint: Consider the automorphism $\psi_\alpha(z) = (z-\alpha)/(1-\overline{\alpha}z)$ and apply the Schwarz lemma to $\psi_{f(w)} \circ f \circ \psi_w^{-1}$.]

(b) Prove that

$$\frac{|f'(z)|}{1-|f(z)|^2} \leq \frac{1}{1-|z|^2} \qquad \text{for all } z \in \mathbb{D}.$$

This result is called the Schwarz-Pick lemma. See Problem 3 for an important application of this lemma.

14. Prove that all conformal mappings from the upper half-plane $\mathbb{H}$ to the unit disc $\mathbb{D}$ take the form

$$e^{i\theta}\frac{z-\beta}{z-\overline{\beta}}, \qquad \theta \in \mathbb{R} \text{ and } \beta \in \mathbb{H}.$$

15. Here are two properties enjoyed by automorphisms of the upper half-plane.

(a) Suppose Φ is an automorphism of $\mathbb{H}$ that fixes three distinct points on the real axis. Then Φ is the identity.

(b) Suppose (x_1, x_2, x_3) and (y_1, y_2, y_3) are two pairs of three distinct points on the real axis with

$$x_1 < x_2 < x_3 \qquad \text{and} \qquad y_1 < y_2 < y_3.$$

Prove that there exists (a unique) automorphism Φ of $\mathbb{H}$ so that $\Phi(x_j) = y_j$, $j = 1, 2, 3$. The same conclusion holds if $y_3 < y_1 < y_2$ or $y_2 < y_3 < y_1$.

16. Let

$$f(z) = \frac{i-z}{i+z} \qquad \text{and} \qquad f^{-1}(w) = i\frac{1-w}{1+w}.$$

(a) Given $\theta \in \mathbb{R}$, find real numbers a, b, c, d such that $ad - bc = 1$, and so that for any $z \in \mathbb{H}$

$$\frac{az+b}{cz+d} = f^{-1}\left(e^{i\theta} f(z)\right).$$

(b) Given $\alpha \in \mathbb{D}$ find real numbers a, b, c, d so that $ad - bc = 1$, and so that for any $z \in \mathbb{H}$

$$\frac{az+b}{cz+d} = f^{-1}\left(\psi_\alpha(f(z))\right),$$

with ψ_α defined in Section 2.1.

(c) Prove that if g is an automorphism of the unit disc, then there exist real numbers a, b, c, d such that $ad - bc = 1$ and so that for any $z \in \mathbb{H}$

$$\frac{az+b}{cz+d} = f^{-1} \circ g \circ f(z).$$

[Hint: Use parts (a) and (b).]

17. If $\psi_\alpha(z) = (\alpha - z)/(1 - \overline{\alpha} z)$ for $|\alpha| < 1$, prove that

$$\frac{1}{\pi} \iint_{\mathbb{D}} |\psi_\alpha'|^2 \, dxdy = 1 \qquad \text{and} \qquad \frac{1}{\pi} \iint_{\mathbb{D}} |\psi_\alpha'| \, dxdy = \frac{1 - |\alpha|^2}{|\alpha|^2} \log \frac{1}{1 - |\alpha|^2},$$

where in the case $\alpha = 0$ the expression on the right is understood as the limit as $|\alpha| \to 0$.

[Hint: The first integral can be evaluated without a calculation. For the second, use polar coordinates, and for each fixed r use contour integration to evaluate the integral in θ.]

18. Suppose that Ω is a simply connected domain that is bounded by a piecewise-smooth closed curve γ (in the terminology of Chapter 1). Then any conformal map F of $\mathbb{D}$ to Ω extends to a continuous bijection of $\overline{\mathbb{D}}$ to $\overline{\Omega}$. The proof is simply a generalization of the argument used in Theorem 4.2.

19. Prove that the complex plane slit along the union of the rays $\cup_{k=1}^{n} \{A_k + iy : y \leq 0\}$ is simply connected.

[Hint: Given a curve, first "raise" it so that it is completely contained in the upper half-plane.]

20. Other examples of elliptic integrals providing conformal maps from the upper half-plane to rectangles are given below.

(a) The function

$$\int_0^z \frac{d\zeta}{\sqrt{\zeta(\zeta-1)(\zeta-\lambda)}}, \qquad \text{with } \lambda \in \mathbb{R} \text{ and } \lambda \neq 1$$

maps the upper half-plane conformally to a rectangle, one of whose vertices is the image of the point at infinity.

(b) In the case $\lambda = -1$, the image of

$$\int_0^z \frac{d\zeta}{\sqrt{\zeta(\zeta^2-1)}}$$

is a square whose side lengths are $\frac{\Gamma^2(1/4)}{2\sqrt{2\pi}}$.

21. We consider conformal mappings to triangles.

(a) Show that

$$\int_0^z z^{-\beta_1}(1-z)^{-\beta_2}\, dz,$$

with $0 < \beta_1 < 1$, $0 < \beta_2 < 1$, and $1 < \beta_1 + \beta_2 < 2$, maps $\mathbb{H}$ to a triangle whose vertices are the images of 0, 1, and ∞, and with angles $\alpha_1\pi$, $\alpha_2\pi$, and $\alpha_3\pi$, where $\alpha_j + \beta_j = 1$ and $\beta_1 + \beta_2 + \beta_3 = 2$.

(b) What happens when $\beta_1 + \beta_2 = 1$?

(c) What happens when $0 < \beta_1 + \beta_2 < 1$?

(d) In (a), the length of the side of the triangle opposite angle $\alpha_j\pi$ is $\frac{\sin(\alpha_j\pi)}{\pi}\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)$.

22. If P is a simply connected region bounded by a polygon with vertices $a_1, \ldots, a_n$ and angles $\alpha_1\pi, \ldots, \alpha_n\pi$, and F is a conformal map of the disc $\mathbb{D}$ to P, then there exist complex numbers $B_1, \ldots, B_n$ on the unit circle, and constants c_1 and c_2 so that

$$F(z) = c_1 \int_1^z \frac{d\zeta}{(\zeta - B_1)^{\beta_1}\cdots(\zeta - B_n)^{\beta_n}} + c_2.$$

[Hint: This follows from the standard correspondence between $\mathbb{H}$ and $\mathbb{D}$ and an argument similar to that used in the proof of Theorem 4.7.]

23. If

$$F(z) = \int_1^z \frac{d\zeta}{(1-\zeta^n)^{2/n}},$$

then F maps the unit disc conformally onto the interior of a regular polygon with n sides and perimeter

$$2^{\frac{n-2}{n}} \int_0^{\pi} (\sin\theta)^{-2/n}\, d\theta.$$

24. The elliptic integrals K and K' defined for $0 < k < 1$ by

$$K(k) = \int_0^1 \frac{dx}{((1-x^2)(1-k^2x^2))^{1/2}} \quad \text{and} \quad K'(k) = \int_1^{1/k} \frac{dx}{((x^2-1)(1-k^2x^2))^{1/2}}$$

satisfy various interesting identities. For instance:

(a) Show that if $\tilde{k}^2 = 1 - k^2$ and $0 < \tilde{k} < 1$, then

$$K'(k) = K(\tilde{k}).$$

[Hint: Change variables $x = (1 - \tilde{k}^2 y^2)^{-1/2}$ in the integral defining $K'(k)$.]

(b) Prove that if $\tilde{k}^2 = 1 - k^2$, and $0 < \tilde{k} < 1$, then

$$K(k) = \frac{2}{1+\tilde{k}}\, K\left(\frac{1-\tilde{k}}{1+\tilde{k}}\right).$$

[Hint: Change variables $x = 2t/(1 + \tilde{k} + (1 - \tilde{k})t^2)$.]

(c) Show that for $0 < k < 1$ one has

$$K(k) = \frac{\pi}{2}\, F(1/2, 1/2, 1; k^2),$$

where F the hypergeometric series. [Hint: This follows from the integral representation for F given in Exercise 9, Chapter 6.]

6 Problems

1. Let f be a complex-valued C^1 function defined in the neighborhood of a point z_0. There are several notions closely related to conformality at z_0. We say that f is **isogonal** at z_0 if whenever $\gamma(t)$ and $\eta(t)$ are two smooth curves with $\gamma(0) = \eta(0) = z_0$, that make an angle θ there ($|\theta| < \pi$), then $f(\gamma(t))$ and $f(\eta(t))$ make an angle of θ' at $t = 0$ with $|\theta'| = |\theta|$ for all θ. Also, f is said to be **isotropic** if it magnifies lengths by some factor for all directions emanating from z_0, that is, if the limit

$$\lim_{r\to 0} \frac{|f(z_0 + re^{i\theta}) - f(z_0)|}{r}$$

exists, is non-zero, and independent of θ.

Then f is isogonal at z_0 if and only if it is isotropic at z_0; moreover, f is isogonal at z_0 if and only if either $f'(z_0)$ exists and is non-zero, or the same holds for f replaced by $\overline{f}$.

2. The angle between two non-zero complex numbers z and w (taken in that order) is simply the oriented angle, in $(-\pi, \pi]$, that is formed between the two vectors in $\mathbb{R}^2$ corresponding to the points z and w. This oriented angle, say α, is uniquely determined by the two quantities

$$\frac{(z,w)}{|z|\,|w|} \quad \text{and} \quad \frac{(z,-iw)}{|z|\,|w|}$$

which are simply the cosine and sine of α, respectively. Here, the notation $(\cdot,\cdot)$ corresponds to the usual Euclidian inner product in $\mathbb{R}^2$, which in terms of complex numbers takes the form $(z,w) = \mathrm{Re}(z\overline{w})$.

In particular, we may now consider two smooth curves $\gamma : [a,b] \to \mathbb{C}$ and $\eta : [a,b] \to \mathbb{C}$, that intersect at z_0, say $\gamma(t_0) = \eta(t_0) = z_0$, for some $t_0 \in (a,b)$. If the quantities $\gamma'(t_0)$ and $\eta'(t_0)$ are non-zero, then they represent the tangents to the curves γ and η at the point z_0, and we say that the two curves intersect at z_0 at the angle formed by the two vectors $\gamma'(t_0)$ and $\eta'(t_0)$.

A holomorphic function f defined near z_0 is said to **preserve angles** at z_0 if for any two smooth curves γ and η intersecting at z_0, the angle formed between the curves γ and η at z_0 equals the angle formed between the curves $f \circ \gamma$ and $f \circ \eta$ at $f(z_0)$. (See Figure 12 for an illustration.) In particular, we assume that the tangents to the curves γ, η, $f \circ \gamma$, and $f \circ \eta$ at the point z_0 and $f(z_0)$ are all non-zero.

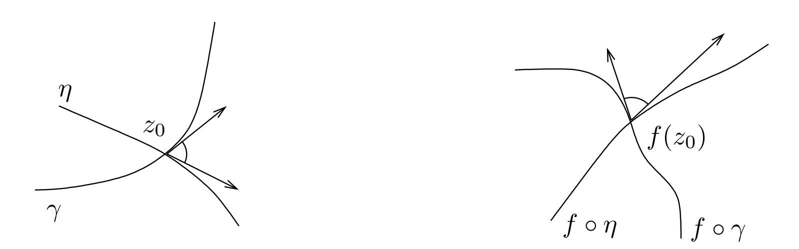


Figure 12. Preservation of angles at z_0

(a) Prove that if $f : \Omega \to \mathbb{C}$ is holomorphic, and $f'(z_0) \neq 0$, then f preserves angles at z_0. [Hint: Observe that

$$(f'(z_0)\gamma'(t_0), f'(z_0)\eta'(t_0)) = |f'(z_0)|^2(\gamma'(t_0), \eta'(t_0)).]$$

(b) Conversely, prove the following: suppose $f : \Omega \to \mathbb{C}$ is a complex-valued function, that is real-differentiable at $z_0 \in \Omega$, and $J_f(z_0) \neq 0$. If f preserves angles at z_0, then f is holomorphic at z_0 with $f'(z_0) \neq 0$.

3.* The Schwarz-Pick lemma (see Exercise 13) is the infinitesimal version of an important observation in complex analysis and geometry.

For complex numbers $w \in \mathbb{C}$ and $z \in \mathbb{D}$ we define the **hyperbolic length** of w at z by

$$\|w\|_z = \frac{|w|}{1-|z|^2},$$

where $|w|$ and $|z|$ denote the usual absolute values. This length is sometimes referred to as the **Poincaré metric**, and as a Riemann metric it is written as

$$ds^2 = \frac{|dz|^2}{(1-|z|^2)^2}.$$

The idea is to think of w as a vector lying in the tangent space at z. Observe that for a fixed w, its hyperbolic length grows to infinity as z approaches the boundary of the disc. We pass from the infinitesimal hyperbolic length of tangent vectors to the global hyperbolic distance between two points by integration.

(a) Given two complex numbers z_1 and z_2 in the disc, we define the **hyperbolic distance** between them by

$$d(z_1, z_2) = \inf_\gamma \int_0^1 \|\gamma'(t)\|_{\gamma(t)}\, dt,$$

where the infimum is taken over all smooth curves $\gamma : [0,1] \to \mathbb{D}$ joining z_1 and z_2. Use the Schwarz-Pick lemma to prove that if $f : \mathbb{D} \to \mathbb{D}$ is holomorphic, then

$$d(f(z_1), f(z_2)) \le d(z_1, z_2) \qquad \text{for any } z_1, z_2 \in \mathbb{D}.$$

In other words, holomorphic functions are distance-decreasing in the hyperbolic metric.

(b) Prove that automorphisms of the unit disc preserve the hyperbolic distance, namely

$$d(\varphi(z_1), \varphi(z_2)) = d(z_1, z_2), \qquad \text{for any } z_1, z_2 \in \mathbb{D}$$

and any automorphism φ. Conversely, if $\varphi : \mathbb{D} \to \mathbb{D}$ preserves the hyperbolic distance, then either φ or $\overline{\varphi}$ is an automorphism of $\mathbb{D}$.

(c) Given two points $z_1, z_2 \in \mathbb{D}$, show that there exists an automorphism φ such that $\varphi(z_1) = 0$ and $\varphi(z_2) = s$ for some s on the segment $[0,1)$ on the real line.

(d) Prove that the hyperbolic distance between 0 and $s \in [0,1)$ is

$$d(0, s) = \frac{1}{2}\log\frac{1+s}{1-s}.$$

(e) Find a formula for the hyperbolic distance between any two points in the unit disc.

4.* Consider the group of matrices of the form

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

that satisfy the following conditions:

(i) a, b, c, and $d \in \mathbb{C}$,

(ii) the determinant of M is equal to 1,

(iii) the matrix M preserves the following hermitian form on $\mathbb{C}^2 \times \mathbb{C}^2$:

$$\langle Z, W \rangle = z_1 \overline{w}_1 - z_2 \overline{w}_2,$$

where $Z = (z_1, z_2)$ and $W = (w_1, w_2)$. In other words, for all $Z, W \in \mathbb{C}^2$

$$\langle MZ, MW \rangle = \langle Z, W \rangle.$$

This group of matrices is denoted by $\mathrm{SU}(1,1)$.

(a) Prove that all matrices in $\mathrm{SU}(1,1)$ are of the form

$$\begin{pmatrix} a & b \\ \overline{b} & \overline{a} \end{pmatrix},$$

where $|a|^2 - |b|^2 = 1$. To do so, consider the matrix

$$J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and observe that $\langle Z, W \rangle = {}^t\overline{W} J Z$, where ${}^t\overline{W}$ denotes the conjugate transpose of W.

(b) To every matrix in $\mathrm{SU}(1,1)$ we can associate a fractional linear transformation

$$\frac{az + b}{cz + d}.$$

Prove that the group $\mathrm{SU}(1,1)/\{\pm 1\}$ is isomorphic to the group of automorphisms of the disc. [Hint: Use the following association.]

$$e^{2i\theta} \frac{z - \alpha}{1 - \overline{\alpha} z} \longrightarrow \begin{pmatrix} \frac{e^{i\theta}}{\sqrt{1-|\alpha|^2}} & -\frac{\alpha e^{i\theta}}{\sqrt{1-|\alpha|^2}} \\ -\frac{\overline{\alpha} e^{-i\theta}}{\sqrt{1-|\alpha|^2}} & \frac{e^{-i\theta}}{\sqrt{1-|\alpha|^2}} \end{pmatrix}.$$

5. The following result is relevant to Problem 4 in Chapter 10 which treats modular functions.

(a) Suppose that $F : \mathbb{H} \to \mathbb{C}$ is holomorphic and bounded. Also, suppose that $F(z)$ vanishes when $z = ir_n$, $n = 1, 2, 3, \ldots$, where $\{r_n\}$ is a bounded sequence of positive numbers. Prove that if $\sum_{n=1}^{\infty} r_n = \infty$, then $F = 0$.

(b) If $\sum r_n < \infty$, it is possible to construct a bounded function on the upper half-plane with zeros precisely at the points ir_n.

For related results in the unit disc, see Problems 1 and 2 in Chapter 5.]

6.* The results of Exercise 18 extend to the case when γ is assumed merely to be closed, simple, and continuous. The proof, however, requires further ideas.

7.* Applying ideas of Carathéodory, Koebe gave a proof of the Riemann mapping theorem by constructing (more explicitly) a sequence of functions that converges to the desired conformal map.

Starting with a Koebe domain, that is, a simply connected domain $\mathcal{K}_0 \subset \mathbb{D}$ that is not all of $\mathbb{D}$, and which contains the origin, the strategy is to find an injective function f_0 such that $f_0(\mathcal{K}_0) = \mathcal{K}_1$ is a Koebe domain "larger" than $\mathcal{K}_0$. Then, one iterates this process, finally obtaining functions $F_n = f_n \circ \cdots \circ f_0 : \mathcal{K}_0 \to \mathbb{D}$ such that $F_n(\mathcal{K}_0) = \mathcal{K}_{n+1}$ and $\lim F_n = F$ is a conformal map from $\mathcal{K}_0$ to $\mathbb{D}$.

The **inner radius** of a region $\mathcal{K} \subset \mathbb{D}$ that contains the origin is defined by $r_{\mathcal{K}} = \sup\{\rho \geq 0 : D(0, \rho) \subset \mathcal{K}\}$. Also, a holomorphic injection $f : \mathcal{K} \to \mathbb{D}$ is said to be an **expansion** if $f(0) = 0$ and $|f(z)| > |z|$ for all $z \in \mathcal{K} - \{0\}$.

(a) Prove that if f is an expansion, then $r_{f(\mathcal{K})} \geq r_{\mathcal{K}}$ and $|f'(0)| > 1$. [Hint: Write $f(z) = zg(z)$ and use the maximum principle to prove that $|f'(0)| = |g(0)| > 1$.]

Suppose we begin with a Koebe domain $\mathcal{K}_0$ and a sequence of expansions $\{f_0, f_1, \ldots, f_n, \ldots\}$, so that $\mathcal{K}_{n+1} = f_n(\mathcal{K}_n)$ are also Koebe domains. We then define holomorphic maps $F_n : \mathcal{K}_0 \to \mathbb{D}$ by $F_n = f_n \circ \cdots \circ f_0$.

(b) Prove that for each n, the function F_n is an expansion. Moreover, $F_n'(0) = \prod_{k=0}^{n} f_k'(0)$, and conclude that $\lim_{n\to\infty} |f_n'(0)| = 1$. [Hint: Prove that the sequence $\{|F_n'(0)|\}$ has a limit by showing that it is bounded above and monotone increasing. Use the Schwarz lemma.]

(c) Show that if the sequence is osculating, that is, $r_{\mathcal{K}_n} \to 1$ as $n \to \infty$, then $\{F_n\}$ converges uniformly on compact subsets of $\mathcal{K}_0$ to a conformal map $F : \mathcal{K}_0 \to \mathbb{D}$. [Hint: If $r_{F(\mathcal{K}_0)} \geq 1$ then F is surjective.]

To construct the desired osculating sequence we shall use the automorphisms $\psi_\alpha = (\alpha - z)/(1 - \overline{\alpha}z)$.

(d) Given a Koebe domain $\mathcal{K}$, choose a point $\alpha \in \mathbb{D}$ on the boundary of $\mathcal{K}$ such that $|\alpha| = r_{\mathcal{K}}$, and also choose $\beta \in \mathbb{D}$ such that $\beta^2 = \alpha$. Let S denote the square root of ψ_α on $\mathcal{K}$ such that $S(0) = 0$. Why is such a function well defined? Prove that the function $f : \mathcal{K} \to \mathbb{D}$ defined by $f(z) = \psi_\beta \circ S \circ \psi_\alpha$

is an expansion. Moreover, show that $|f'(0)| = (1 + r_{\mathcal{K}})/2\sqrt{r_{\mathcal{K}}}$. [Hint: To prove that $|f(z)| > |z|$ on $\mathcal{K} - \{0\}$ apply the Schwarz lemma to the inverse function, namely $\psi_\alpha \circ g \circ \psi_\beta$ where $g(z) = z^2$.]

(e) Use part (d) to construct the desired sequence.

8.* Let f be an injective holomorphic function in the unit disc, with $f(0) = 0$ and $f'(0) = 1$. If we write $f(z) = z + a_2z^2 + a_3z^3 \cdots$, then Problem 1 in Chapter 3 shows that $|a_2| \leq 2$. Bieberbach conjectured that in fact $|a_n| \leq n$ for all $n \geq 2$; this was proved by deBranges. This problem outlines an argument to prove the conjecture under the additional assumption that the coefficients a_n are real.

(a) Let $z = re^{i\theta}$ with $0 < r < 1$, and show that if $v(r, \theta)$ denotes the imaginary part of $f(re^{i\theta})$, then

$$a_n r^n = \frac{2}{\pi} \int_0^\pi v(r, \theta) \sin n\theta \, d\theta.$$

(b) Show that for $0 \leq \theta \leq \pi$ and $n = 1, 2, \ldots$ we have $|\sin n\theta| \leq n \sin \theta$.

(c) Use the fact that $a_n \in \mathbb{R}$ to show that $f(\mathbb{D})$ is symmetric with respect to the real axis, and use this fact to show that f maps the upper half-disc into either the upper or lower part of $f(\mathbb{D})$.

(d) Show that for r small,

$$v(r, \theta) = r \sin \theta [1 + O(r)],$$

and use the previous part to conclude that $v(r, \theta) \sin \theta \geq 0$ for all $0 < r < 1$ and $0 \leq \theta \leq \pi$.

(e) Prove that $|a_n r^n| \leq nr$, and let $r \to 1$ to conclude that $|a_n| \leq n$.

(f) Check that the function $f(z) = z/(1 - z)^2$ satisfies all the hypotheses and that $|a_n| = n$ for all n.

9.* Gauss found a connection between elliptic integrals and the familiar operations of forming arithmetic and geometric means.

We start with any pair (a, b) of numbers that satisfy $a \geq b > 0$, and form the arithmetic and geometric means of a and b, that is,

$$a_1 = \frac{a + b}{2} \quad \text{and} \quad b_1 = (ab)^{1/2}.$$

We then repeat these operations with a and b replaced by a_1 and b_1. Iterating this process provides two sequences $\{a_n\}$ and $\{b_n\}$ where a_{n+1} and b_{n+1} are the arithmetic and geometric means of a_n and b_n, respectively.

(a) Prove that the two sequences $\{a_n\}$ and $\{b_n\}$ have a common limit. This limit, which we denote by $M(a,b)$, is called the **arithmetic-geometric mean** of a and b. [Hint: Show that $a \geq a_1 \geq a_2 \geq \cdots \geq a_n \geq b_n \geq \cdots \geq b_1 \geq b$ and $a_n - b_n \leq (a-b)/2^n$.]

(b) Gauss's identity states that

$$\frac{1}{M(a,b)} = \frac{2}{\pi}\int_0^{\pi/2} \frac{d\theta}{(a^2\cos^2\theta + b^2\sin^2\theta)^{1/2}}.$$

To prove this relation, show that if $I(a,b)$ denotes the integral on the right-hand side, then it suffices to establish the invariance of I, namely

$$I(a,b) = I\left(\frac{a+b}{2}, (ab)^{1/2}\right). \tag{6}$$

Then, observe that the connection with elliptic integrals takes the form

$$I(a,b) = \frac{1}{a}K(k) = \frac{1}{a}\int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \qquad \text{where } k^2 = 1 - b^2/a^2,$$

and that the relation (6) is a consequence of the identity in (b) of Exercise 24.

9 An Introduction to Elliptic Functions

> The form that Jacobi had given to the theory of elliptic functions was far from perfection; its flaws are obvious. At the base we find three fundamental functions sn, cn and dn. These functions do not have the same periods...
>
> In Weierstrass' system, instead of three fundamental functions, there is only one, $\wp(u)$, and it is the simplest of all having the same periods. It has only one double infinity; and finally its definition is so that it does not change when one replaces one system of periods by another equivalent system.
>
> *H. Poincaré,* 1899

The theory of elliptic functions, which is of interest in several parts of mathematics, initially grew out of the study of elliptic integrals. These can be described generally as integrals of the form $\int R(x, \sqrt{P(x)})\, dx$, where R is a rational function and P a polynomial of degree three or four.[1] These integrals arose in computing the arc-length of an ellipse, or of a lemniscate, and in a variety of other problems. Their early study was centered on their special transformation properties and on the discovery of an inherent double-periodicity. We have seen an example of this latter phenomenon in the mapping function of the half-plane to a rectangle taken up in Section 4.5 of the previous chapter.

It was Jacobi who transformed the subject by initiating the systematic study of doubly-periodic functions (called elliptic functions). In this theory, the theta functions he introduced played a decisive role. Weierstrass after him developed another approach, which in its initial steps is simpler and more elegant. It is based on his $\wp$ function, and in this chapter we shall sketch the beginnings of that theory. We will go as far as to glimpse a possible connection with number theory, by considering the Eisenstein series and their expression involving divisor functions. A number of more direct links with combinatorics and number theory arise from the theta

[1] The case when P is a quadratic polynomial is essentially that of "circular functions", and can be reduced to the trigonometric functions $\sin x$, $\cos x$, etc.

functions, which we will take up in the next chapter. The remarkable facts we shall see there attest to the great interest of these functions in mathematics. As such they ought to soften the harsh opinion expressed above about the imperfection of Jacobi's theory.

1 Elliptic functions

We are interested in meromorphic functions f on $\mathbb{C}$ that have two periods; that is, there are two non-zero complex numbers ω_1 and ω_2 such that

$$f(z+\omega_1) = f(z) \quad \text{and} \quad f(z+\omega_2) = f(z),$$

for all $z \in \mathbb{C}$. A function with two periods is said to be **doubly periodic**.

The case when ω_1 and ω_2 are linearly dependent over $\mathbb{R}$, that is $\omega_2/\omega_1 \in \mathbb{R}$, is uninteresting. Indeed, Exercise 1 shows that in this case f is either periodic with a simple period (if the quotient ω_2/ω_1 is rational) or f is constant (if ω_2/ω_1 is irrational). Therefore, we make the following assumption: the periods ω_1 and ω_2 are linearly independent over $\mathbb{R}$.

We now describe a normalization that we shall use extensively in this chapter. Let $\tau = \omega_2/\omega_1$. Since τ and $1/\tau$ have imaginary parts of opposite signs, and since τ is not real, we may assume (after possibly interchanging the roles of ω_1 and ω_2) that $\mathrm{Im}(\tau) > 0$. Observe now that the function f has periods ω_1 and ω_2 if and only if the function $F(z) = f(\omega_1 z)$ has periods 1 and τ, and moreover, the function f is meromorphic if and only if F is meromorphic. Also the properties of f are immediately deducible from those of F. We may therefore assume, without loss of generality, that f is a meromorphic function on $\mathbb{C}$ with periods 1 and τ where $\mathrm{Im}(\tau) > 0$.

Successive applications of the periodicity conditions yield

$$f(z+n+m\tau) = f(z) \quad \text{for all integers } n, m \text{ and all } z \in \mathbb{C}, \tag{1}$$

and it is therefore natural to consider the lattice in $\mathbb{C}$ defined by

$$\Lambda = \{n + m\tau : \; n, m \in \mathbb{Z}\}.$$

We say that 1 and τ **generate** Λ (see Figure 1).

Equation (1) says that f is constant under translations by elements of Λ. Associated to the lattice Λ is the **fundamental parallelogram** defined by

$$P_0 = \{z \in \mathbb{C} : z = a + b\tau \;\; \text{where } 0 \le a < 1 \text{ and } 0 \le b < 1\}.$$

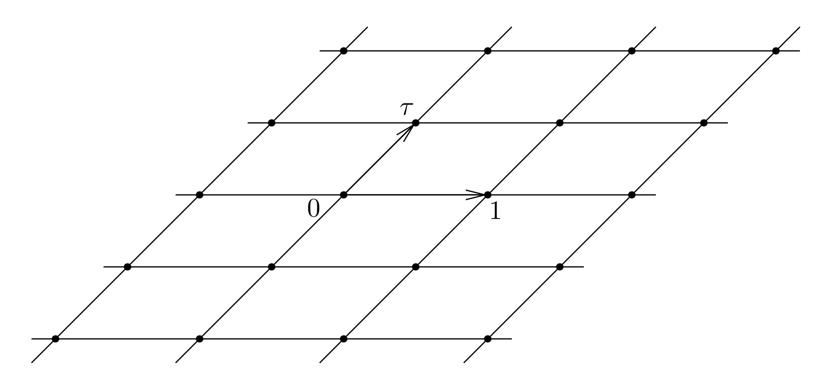


Figure 1. The lattice Λ generated by 1 and τ

The importance of the fundamental parallelogram comes from the fact that f is completely determined by its behavior on P_0. To see this, we need a definition: two complex numbers z and w are **congruent modulo** Λ if

$$z = w + n + m\tau \qquad \text{for some } n, m \in \mathbb{Z},$$

and we write $z \sim w$. In other words, z and w differ by a point in the lattice, $z - w \in \Lambda$. By (1) we conclude that $f(z) = f(w)$ whenever $z \sim w$. If we can show that any point in $z \in \mathbb{C}$ is congruent to a unique point in P_0 then we will have proved that f is completely determined by its values in the fundamental parallelogram. Suppose $z = x + iy$ is given, and write $z = a + b\tau$ where $a, b \in \mathbb{R}$. This is possible since 1 and τ form a basis over the reals of the two-dimensional vector space $\mathbb{C}$. Then choose n and m to be the greatest integers $\leq a$ and $\leq b$, respectively. If we let $w = z - n - m\tau$, then by definition $z \sim w$, and moreover $w = (a - n) + (b - m)\tau$. By construction, it is clear that $w \in P_0$. To prove uniqueness, suppose that w and w' are two points in P_0 that are congruent. If we write $w = a + b\tau$ and $w' = a' + b'\tau$, then $w - w' = (a - a') + (b - b')\tau \in \Lambda$, and therefore both $a - a'$ and $b - b'$ are integers. But since $0 \leq a, a' < 1$, we have $-1 < a - a' < 1$, which then implies $a - a' = 0$. Similarly $b - b' = 0$, and we conclude that $w = w'$.

More generally, a **period parallelogram** P is any translate of the fundamental parallelogram, $P = P_0 + h$ with $h \in \mathbb{C}$ (see Figure 2).

Since we can apply the lemma to $z - h$, we conclude that every point in $\mathbb{C}$ is congruent to a unique point in a given period parallelogram. Therefore, f is uniquely determined by its behavior on any period parallelogram.

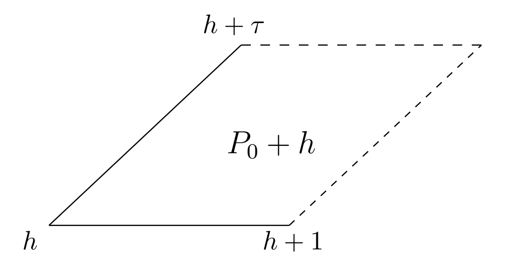


Figure 2. A period parallelogram

Finally, note that Λ and P_0 give rise to a covering (or tiling) of the complex plane

$$\mathbb{C} = \bigcup_{n,m\in\mathbb{Z}} (n + m\tau + P_0), \tag{2}$$

and moreover, this union is disjoint. This is immediate from the facts we just collected and the definition of P_0. We summarize what we have seen so far.

Proposition 1.1 *Suppose f is a meromorphic function with two periods 1 and τ which generate the lattice Λ. Then:*

(i) *Every point in $\mathbb{C}$ is congruent to a unique point in the fundamental parallelogram.*

(ii) *Every point in $\mathbb{C}$ is congruent to a unique point in any given period parallelogram.*

(iii) *The lattice Λ provides a disjoint covering of the complex plane, in the sense of* (2).

(iv) *The function f is completely determined by its values in any period parallelogram.*

1.1 Liouville's theorems

We can now see why we assumed from the beginning that f is meromorphic rather than just holomorphic.

Theorem 1.2 *An entire doubly periodic function is constant.*

Proof. The function is completely determined by its values on P_0 and since the closure of P_0 is compact, we conclude that the function is bounded on $\mathbb{C}$, hence constant by Liouville's theorem in Chapter 2.

A non-constant doubly periodic meromorphic function is called an **elliptic function**. Since a meromorphic function can have only finitely many zeros and poles in any large disc, we see that an elliptic function will have only finitely many zeros and poles in any given period parallelogram, and in particular, this is true in the fundamental parallelogram. Of course, nothing excludes f from having a pole or zero on the boundary of P_0.

As usual, we count poles and zeros with multiplicities. Keeping this in mind we can prove the following theorem.

Theorem 1.3 *The total number of poles of an elliptic function in P_0 is always ≥ 2.*

In other words, f cannot have only one simple pole. It must have at least two poles, and this does not exclude the case of a single pole of multiplicity ≥ 2.

Proof. Suppose first that f has no poles on the boundary ∂P_0 of the fundamental parallelogram. By the residue theorem we have

$$\int_{\partial P_0} f(z)\, dz = 2\pi i \sum \operatorname{res} f,$$

and we contend that the integral is 0. To see this, we simply use the periodicity of f. Note that

$$\int_{\partial P_0} f(z)\, dz = \int_0^1 f(z)\, dz + \int_1^{1+\tau} f(z)\, dz + \int_{1+\tau}^{\tau} f(z)\, dz + \int_{\tau}^{0} f(z)\, dz,$$

and the integrals over opposite sides cancel out. For instance

$$\begin{aligned}
\int_0^1 f(z)\, dz + \int_{1+\tau}^{\tau} f(z)\, dz &= \int_0^1 f(z)\, dz + \int_1^0 f(s+\tau)\, ds \\
&= \int_0^1 f(z)\, dz + \int_1^0 f(s)\, ds \\
&= \int_0^1 f(z)\, dz - \int_0^1 f(z)\, dz \\
&= 0,
\end{aligned}$$

and similarly for the other pair of sides. Hence $\int_{\partial P_0} f = 0$ and $\sum \operatorname{res} f = 0$. Therefore f must have at least two poles in P_0.

If f has a pole on ∂P_0 choose a small $h \in \mathbb{C}$ so that if $P = h + P_0$, then f has no poles on ∂P. Arguing as before, we find that f must have at least two poles in P, and therefore the same conclusion holds for P_0.

The total number of poles (counted according to their multiplicities) of an elliptic function is called its **order**. The next theorem says that elliptic functions have as many zeros as they have poles, if the zeros are counted with their multiplicities.

Theorem 1.4 *Every elliptic function of order m has m zeros in P_0.*

Proof. Assuming first that f has no zeros or poles on the boundary of P_0, we know by the argument principle in Chapter 3 that

$$\int_{\partial P_0} \frac{f'(z)}{f(z)}\, dz = 2\pi i(\mathcal{N}_{\mathfrak{z}} - \mathcal{N}_{\mathfrak{p}})$$

where $\mathcal{N}_{\mathfrak{z}}$ and $\mathcal{N}_{\mathfrak{p}}$ denote the number of zeros and poles of f in P_0, respectively. By periodicity, we can argue as in the proof of the previous theorem to find that $\int_{\partial P_0} f'/f = 0$, and therefore $\mathcal{N}_{\mathfrak{z}} = \mathcal{N}_{\mathfrak{p}}$.

In the case when a pole or zero of f lies on ∂P_0 it suffices to apply the argument to a translate of P.

As a consequence, if f is elliptic then the equation $f(z) = c$ has as many solutions as the order of f for every $c \in \mathbb{C}$, simply because $f - c$ is elliptic and has as many poles as f.

Despite the rather simple nature of the theorems above, there remains the question of showing that elliptic functions exist. We now turn to a constructive solution of this problem.

1.2 The Weierstrass $\wp$ function

An elliptic function of order two

This section is devoted to the basic example of an elliptic function. As we have seen above, any elliptic function must have at least two poles; we shall in fact construct one whose only singularity will be a double pole at the points of the lattice generated by the periods.

Before looking at the case of doubly-periodic functions, let us first consider briefly functions with only a single period. If one wished to construct a function with period 1 and poles at all the integers, a simple choice would be the sum

$$F(z) = \sum_{n=-\infty}^{\infty} \frac{1}{z+n}.$$

Note that the sum remains unchanged if we replace z by $z + 1$, and the poles are at the integers. However, the series defining F is not absolutely

convergent, and to remedy this problem, we sum symmetrically, that is, we define

$$F(z) = \lim_{N\to\infty} \sum_{|n|\leq N} \frac{1}{z+n} = \frac{1}{z} + \sum_{n=1}^{\infty} \left[\frac{1}{z+n} + \frac{1}{z-n} \right].$$

On the far right-hand side, we have paired up the terms corresponding to n and $-n$, a trick which makes the quantity in brackets $O(1/n^2)$, and hence the last sum is absolutely convergent. As a consequence, F is meromorphic with poles precisely at the integers. In fact, we proved earlier in Chapter 5 that $F(z) = \pi \cot \pi z$.

There is a second way to deal with the series $\sum_{-\infty}^{\infty} 1/(z+n)$, which is to write it as

$$\frac{1}{z} + \sum_{n\neq 0} \left[\frac{1}{z+n} - \frac{1}{n} \right],$$

where the sum is taken over all non-zero integers. Notice that $1/(z+n) - 1/n = O(1/n^2)$, which makes this series absolutely convergent. Moreover, since

$$\frac{1}{z+n} + \frac{1}{z-n} = \left(\frac{1}{z+n} - \frac{1}{n} \right) + \left(\frac{1}{z-n} - \frac{1}{-n} \right),$$

we get the same sum as before.

In analogy to this, the idea is to mimic the above to produce our first example of an elliptic function. We would like to write it as

$$\sum_{\omega\in\Lambda} \frac{1}{(z+\omega)^2},$$

but again this series does not converge absolutely. There are several approaches to try to make sense of this series (see Problem 1), but the simplest is to follow the second way we dealt with the cotangent series.

To overcome the non-absolute convergence of the series, let Λ^* denote the lattice minus the origin, that is, $\Lambda^* = \Lambda - \{(0,0)\}$, and consider instead the following series:

$$\frac{1}{z^2} + \sum_{\omega\in\Lambda^*} \left[\frac{1}{(z+\omega)^2} - \frac{1}{\omega^2} \right],$$

where we have subtracted the factor $1/\omega^2$ to make the sum converge. The term in brackets is now

$$\frac{1}{(z+\omega)^2} - \frac{1}{\omega^2} = \frac{-z^2 - 2z\omega}{(z+\omega)^2\omega^2} = O\left(\frac{1}{\omega^3}\right) \quad \text{as } |\omega| \to \infty,$$

and the new series will define a meromorphic function with the desired poles once we have proved the following lemma.

Lemma 1.5 *The two series*

$$\sum_{(n,m)\neq(0,0)} \frac{1}{(|n|+|m|)^r} \qquad \text{and} \qquad \sum_{n+m\tau\in\Lambda^*} \frac{1}{|n+m\tau|^r}$$

converge if $r > 2$.

Recall that according to the Note at the end of Chapter 7, the question whether a double series converges absolutely is independent of the order of summation. In the present case, we shall first sum in m and then in n.

For the first series, the usual integral comparison can be applied.[2] For each $n \neq 0$

$$\begin{aligned}\sum_{m\in\mathbb{Z}} \frac{1}{(|n|+|m|)^r} &= \frac{1}{|n|^r} + 2\sum_{m\geq 1} \frac{1}{(|n|+|m|)^r} \\ &= \frac{1}{|n|^r} + 2\sum_{k\geq |n|+1} \frac{1}{k^r} \\ &\leq \frac{1}{|n|^r} + 2\int_{|n|}^{\infty} \frac{dx}{x^r} \\ &\leq \frac{1}{|n|^r} + C\frac{1}{|n|^{r-1}}.\end{aligned}$$

Therefore, $r > 2$ implies

$$\begin{aligned}\sum_{(n,m)\neq(0,0)} \frac{1}{(|n|+|m|)^r} &= \sum_{|m|\neq 0} \frac{1}{|m|^r} + \sum_{|n|\neq 0}\sum_{m\in\mathbb{Z}} \frac{1}{(|n|+|m|)^r} \\ &\leq \sum_{|m|\neq 0} \frac{1}{|m|^r} + \sum_{|n|\neq 0}\left(\frac{1}{|n|^r} + C\frac{1}{|n|^{r-1}}\right) \\ &< \infty.\end{aligned}$$

To prove that the second series also converges, it suffices to show that there is a constant c such that

$$|n| + |m| \leq c|n+\tau m| \qquad \text{for all } n, m \in \mathbb{Z}.$$

---

[2] We simply use $1/k^r \leq 1/x^r$ when $k-1 \leq x \leq k$; see also the first figure in Chapter 8, Book I.

We use the notation $x \lesssim y$ if there exists a positive constant a such that $x \leq ay$. We also write $x \approx y$ if both $x \lesssim y$ and $y \lesssim x$ hold. Note that for any two positive numbers A and B, one has

$$(A^2 + B^2)^{1/2} \approx A + B.$$

On the one hand $A \leq (A^2 + B^2)^{1/2}$ and $B \leq (A^2 + B^2)^{1/2}$, so that $A + B \leq 2(A^2 + B^2)^{1/2}$. On the other hand, it suffices to square both sides to see that $(A^2 + B^2)^{1/2} \leq A + B$.

The proof that the second series in Lemma 1.5 converges is now a consequence of the following observation:

$$|n| + |m| \approx |n + m\tau| \quad \text{whenever } \tau \in \mathbb{H}.$$

Indeed, if $\tau = s + it$ with $s, t \in \mathbb{R}$ and $t > 0$, then

$$|n + m\tau| = [(n + ms)^2 + (mt)^2]^{1/2} \approx |n + ms| + |mt| \approx |n + ms| + |m|,$$

by the previous observation. Then, $|n + ms| + |m| \approx |n| + |m|$, by considering separately the cases when $|n| \leq 2\,|m|\,|s|$ and $|n| \geq 2\,|m|\,|s|$.

Remark. The proof above shows that when $r > 2$ the series $\sum |n + m\tau|^{-r}$ converges uniformly in every half-plane $\mathrm{Im}(\tau) \geq \delta > 0$. In contrast, when $r = 2$ this series fails to converge (Exercise 3).

With this technical point behind us, we may now return to the definition of the **Weierstrass $\wp$ function**, which is given by the series

$$\begin{aligned}
\wp(z) &= \frac{1}{z^2} + \sum_{\omega \in \Lambda^*} \left[\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right] \\
&= \frac{1}{z^2} + \sum_{(n,m) \neq (0,0)} \left[\frac{1}{(z + n + m\tau)^2} - \frac{1}{(n + m\tau)^2} \right].
\end{aligned}$$

We claim that $\wp$ is a meromorphic function with double poles at the lattice points. To see this, suppose that $|z| < R$, and write

$$\wp(z) = \frac{1}{z^2} + \sum_{|\omega| \leq 2R} \left[\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right] + \sum_{|\omega| > 2R} \left[\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right].$$

The term in the second sum is $O(1/|\omega|^3)$ uniformly for $|z| < R$, so by Lemma 1.5 this second sum defines a holomorphic function in $|z| < R$. Finally, note that the first sum exhibits double poles at the lattice points in the disc $|z| < R$.

Observe that because of the insertion of the terms $-1/\omega^2$, it is no longer obvious whether $\wp$ is doubly periodic. Nevertheless this is true, and $\wp$ has all the properties of an elliptic function of order 2. We gather this result in a theorem.

Theorem 1.6 *The function* $\wp$ *is an elliptic function that has periods* 1 *and* τ*, and double poles at the lattice points.*

Proof. It remains only to prove that $\wp$ is periodic with the correct periods. To do so, note that the derivative is given by differentiating the series for $\wp$ termwise so

$$\wp'(z) = -2 \sum_{n,m\in\mathbb{Z}} \frac{1}{(z+n+m\tau)^3}.$$

This accomplishes two things for us. First, the differentiated series converges absolutely whenever z is not a lattice point, by the case $r = 3$ of Lemma 1.5. Second, the differentiation also eliminates the subtraction term $1/\omega^2$; therefore the series for $\wp'$ is clearly periodic with periods 1 and τ, since it remains unchanged after replacing z by $z + 1$ or $z + \tau$.

Hence, there are two constants a and b such that

$$\wp(z+1) = \wp(z) + a \quad \text{and} \quad \wp(z+\tau) = \wp(z) + b.$$

It is clear from the definition, however, that $\wp$ is even, that is, $\wp(z) = \wp(-z)$, since the sum over $\omega \in \Lambda$ can be replaced by the sum over $-\omega \in \Lambda$. Therefore $\wp(-1/2) = \wp(1/2)$ and $\wp(-\tau/2) = \wp(\tau/2)$, and setting $z = -1/2$ and $z = -\tau/2$, respectively, in the two expressions above proves that $a = b = 0$.

A direct proof of the periodicity of $\wp$ can be given without differentiation; see Exercise 4.

Properties of $\wp$

Several remarks are in order. First, we have already observed that $\wp$ is even, and therefore $\wp'$ is odd. Since $\wp'$ is also periodic with periods 1 and τ, we find that

$$\wp'(1/2) = \wp'(\tau/2) = \wp'\left(\frac{1+\tau}{2}\right) = 0.$$

Indeed, one has, for example,

$$\wp'(1/2) = -\wp'(-1/2) = -\wp'(-1/2+1) = -\wp'(1/2).$$

Since $\wp'$ is elliptic and has order 3, the three points $1/2$, $\tau/2$, and $(1+\tau)/2$ (which are called the **half-periods**) are the only roots of $\wp'$ in the fundamental parallelogram, and they have multiplicity 1. Therefore, if we define

$$\wp(1/2) = e_1, \qquad \wp(\tau/2) = e_2, \qquad \text{and} \qquad \wp\left(\frac{1+\tau}{2}\right) = e_3,$$

we conclude that the equation $\wp(z) = e_1$ has a double root at $1/2$. Since $\wp$ has order 2, there are no other solutions to the equation $\wp(z) = e_1$ in the fundamental parallelogram. Similarly the equations $\wp(z) = e_2$ and $\wp(z) = e_3$ have only double roots at $\tau/2$ and $(1+\tau)/2$, respectively. In particular, the three numbers e_1, e_2, and e_3 are distinct, for otherwise $\wp$ would have at least four roots in the fundamental parallelogram, contradicting the fact that $\wp$ has order 2. From these observations we can prove the following theorem.

Theorem 1.7 *The function $(\wp')^2$ is the cubic polynomial in $\wp$*

$$(\wp')^2 = 4(\wp - e_1)(\wp - e_2)(\wp - e_3).$$

Proof. The only roots of $F(z) = (\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3)$ in the fundamental parallelogram have multiplicity 2 and are at the points $1/2, \tau/2$, and $(1+\tau)/2$. Also, $(\wp')^2$ has double roots at these points. Moreover, F has poles of order 6 at the lattice points, and so does $(\wp')^2$ (because $\wp'$ has poles of order 3 there). Consequently $(\wp')^2/F$ is holomorphic and still doubly-periodic, hence this quotient is constant. To find the value of this constant we note that for z near 0, one has

$$\wp(z) = \frac{1}{z^2} + \cdots \qquad \text{and} \qquad \wp'(z) = \frac{-2}{z^3} + \cdots,$$

where the dots indicate terms of higher order. Therefore the constant is 4, and the theorem is proved.

We next demonstrate the universality of $\wp$ by showing that every elliptic function is a simple combination of $\wp$ and $\wp'$.

Theorem 1.8 *Every elliptic function f with periods 1 and τ is a rational function of $\wp$ and $\wp'$.*

The theorem will be an easy consequence of the following version of it.

Lemma 1.9 *Every even elliptic function F with periods 1 and τ is a rational funcion of $\wp$.*

Proof. If F has a zero or pole at the origin it must be of even order, since F is an even function. As a consequence, there exists an integer m so that $F\wp^m$ has no zero or pole at the lattice points. We may therefore assume that F itself has no zero or pole on Λ.

Our immediate goal is to use $\wp$ to construct a doubly-periodic function G with precisely the same zeros and poles as F. To achieve this, we recall that $\wp(z) - \wp(a)$ has a single zero of order 2 if a is a half-period, and two distinct zeros at a and $-a$ otherwise. We must therefore carefully count the zeros and poles of F.

If a is a zero of F, then so is $-a$, since F is even. Moreover, a is congruent to $-a$ if and only if it is a half-period, in which case the zero is of even order. Therefore, if the points $a_1, -a_1, \ldots, a_m, -a_m$ counted with multiplicities[3] describe all the zeros of F, then

$$[\wp(z) - \wp(a_1)] \cdots [\wp(z) - \wp(a_m)]$$

has precisely the same roots as F. A similar argument, where $b_1, -b_1, \ldots, b_m, -b_m$ (with multiplicities) describe all the poles of F, then shows that

$$G(z) = \frac{[\wp(z) - \wp(a_1)] \cdots [\wp(z) - \wp(a_m)]}{[\wp(z) - \wp(b_1)] \cdots [\wp(z) - \wp(b_m)]}$$

is periodic and has the same zeros and poles as F. Therefore, F/G is holomorphic and doubly-periodic, hence constant. This concludes the proof of the lemma.

To prove the theorem, we first recall that $\wp$ is even while $\wp'$ odd. We then write f as a sum of an even and an odd function,

$$f(z) = f_{\text{even}}(z) + f_{\text{odd}}(z),$$

where in fact

$$f_{\text{even}}(z) = \frac{f(z) + f(-z)}{2} \quad \text{and} \quad f_{\text{odd}}(z) = \frac{f(z) - f(-z)}{2}.$$

Then, since $f_{\text{odd}}/\wp'$ is even, it is clear from the lemma applied to f_{even} and $f_{\text{odd}}/\wp'$ that f is a rational function of $\wp$ and $\wp'$.

[3]If a_j is not a half-period, then a_j and $-a_j$ have the multiplicity of F at these points. If a_j is a half-period, then a_j and $-a_j$ are congruent and each has multiplicity half of the multiplicity of F at this point.

2 The modular character of elliptic functions and Eisenstein series

We shall now study the modular character of elliptic functions, that is, their dependence on τ.

Recall the normalization we made at the beginning of the chapter. We started with two periods ω_1 and ω_2 linearly that are independent over $\mathbb{R}$, and we defined $\tau = \omega_2/\omega_1$. We could then assume that $\text{Im}(\tau) > 0$, and also that the two periods are 1 and τ. Next, we considered the lattice generated by 1 and τ and constructed the function $\wp$, which is elliptic of order 2 with periods 1 and τ. Since the construction of $\wp$ depends on τ, we could write $\wp_\tau$ instead. This leads us to change our point of view and think of $\wp_\tau(z)$ primarily as a function of τ. This approach yields many interesting new insights.

Our considerations are guided by the following observations. First, since 1 and τ generate the periods of $\wp_\tau(z)$, and 1 and $\tau + 1$ generate the same periods, we can expect a close relationship between $\wp_\tau(z)$ and $\wp_{\tau+1}(z)$. In fact, it is easy to see that they are identical. Second, since $\tau = \omega_2/\omega_1$, by the normalization imposed at the beginning of Section 1, we see that $-1/\tau = -\omega_1/\omega_2$ (with $\text{Im}(-1/\tau) > 0$). This corresponds essentially to an interchange of the two periods ω_1 and ω_2, and thus we can also expect an intimate connection between $\wp_\tau$ and $\wp_{-1/\tau}$. In fact, it is easy to verify that $\wp_{-1/\tau}(z) = \tau^2\wp_\tau(\tau z)$.

So we are led to consider the group of transformations of the upper half-plane $\text{Im}(\tau) > 0$, generated by the two transformations $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$. This group is called the **modular group**. On the basis of what we said, it can be expected that all quantities intrinsically attached to $\wp_\tau(z)$ reflect the above transformations. We see this clearly when we consider the Eisenstein series.

2.1 Eisenstein series

The **Eisenstein series** of order k is defined by

$$E_k(\tau) = \sum_{(n,m)\neq(0,0)} \frac{1}{(n + m\tau)^k},$$

whenever k is an integer ≥ 3 and τ is a complex number with $\text{Im}(\tau) > 0$. If Λ is the lattice generated by 1 and τ, and if we write $\omega = n + m\tau$, then another expression for the Eisenstein series is $\sum_{\omega\in\Lambda^*} 1/\omega^k$.

Theorem 2.1 *Eisenstein series have the following properties:*

(i) *The series* $E_k(\tau)$ *converges if* $k \geq 3$*, and is holomorphic in the upper half-plane.*

(ii) $E_k(\tau) = 0$ *if* k *is odd.*

(iii) $E_k(\tau)$ *satisfies the following transformation relations:*

$$E_k(\tau + 1) = E_k(\tau) \quad \text{and} \quad E_k(\tau) = \tau^{-k} E_k(-1/\tau).$$

The last property is sometimes referred to as the **modular** character of the Eisenstein series. We shall return to these and other modular identities in the next chapter.

Proof. By Lemma 1.5 and the remark after it, the series $E_k(\tau)$ converges absolutely and uniformly in every half-plane $\text{Im}(\tau) \geq \delta > 0$, whenever $k \geq 3$; hence $E_k(\tau)$ is holomorphic in the upper half-plane $\text{Im}(\tau) > 0$.

By symmetry, replacing n and m by $-n$ and $-m$, we see that whenever k is odd the Eisenstein series is identically zero.

Finally, the fact that $E_k(\tau)$ is periodic of period 1 is clear from the fact that $n + m(\tau + 1) = n + m + m\tau$, and that we can rearrange the sum by replacing $n + m$ by n. Also, we have

$$(n + m(-1/\tau))^k = \tau^{-k}(n\tau - m)^k,$$

and again we can rearrange the sum, this time replacing $(-m, n)$ by (n, m). Conclusion (iii) then follows.

Remark. Because of the second property, some authors define the Eisenstein series of order k to be $\sum_{(n,m)\neq(0,0)} 1/(n + m\tau)^{2k}$, possibly also with a constant factor in front.

The connection of the E_k with the Weierstrass $\wp$ function arises when we investigate the series expansion of $\wp$ near 0.

Theorem 2.2 *For* z *near* 0*, we have*

$$\begin{aligned}\wp(z) &= \frac{1}{z^2} + 3E_4 z^2 + 5E_6 z^4 + \cdots \\ &= \frac{1}{z^2} + \sum_{k=1}^{\infty} (2k+1) E_{2k+2} z^{2k}.\end{aligned}$$

Proof. From the definition of $\wp$, if we note that we may replace ω by $-\omega$ without changing the sum, we have

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda^*} \left[\frac{1}{(z+\omega)^2} - \frac{1}{\omega^2} \right] = \frac{1}{z^2} + \sum_{\omega \in \Lambda^*} \left[\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right],$$

where $\omega = n + m\tau$. The identity

$$\frac{1}{(1-w)^2} = \sum_{\ell=0}^{\infty} (\ell+1) w^\ell, \qquad \text{for } |w| < 1,$$

which follows from differentiating the geometric series, implies that for all small z

$$\frac{1}{(z-\omega)^2} = \frac{1}{\omega^2} \sum_{\ell=0}^{\infty} (\ell+1) \left(\frac{z}{\omega}\right)^\ell = \frac{1}{\omega^2} + \frac{1}{\omega^2} \sum_{\ell=1}^{\infty} (\ell+1) \left(\frac{z}{\omega}\right)^\ell .$$

Therefore

$$\begin{aligned}
\wp(z) &= \frac{1}{z^2} + \sum_{\omega \in \Lambda^*} \sum_{\ell=1}^{\infty} (\ell+1) \frac{z^\ell}{\omega^{\ell+2}} \\
&= \frac{1}{z^2} + \sum_{\ell=1}^{\infty} (\ell+1) \left(\sum_{\omega \in \Lambda^*} \frac{1}{\omega^{\ell+2}} \right) z^\ell \\
&= \frac{1}{z^2} + \sum_{\ell=1}^{\infty} (\ell+1) E_{\ell+2} z^\ell \\
&= \frac{1}{z^2} + \sum_{k=1}^{\infty} (2k+1) E_{2k+2} z^{2k},
\end{aligned}$$

where we have used the fact that $E_{\ell+2} = 0$ whenever ℓ is odd.

From this theorem, we obtain the following three expansions for z near 0:

$$\begin{aligned}
\wp'(z) &= \frac{-2}{z^3} + 6E_4 z + 20 E_6 z^3 + \cdots , \\
(\wp'(z))^2 &= \frac{4}{z^6} - \frac{24 E_4}{z^2} - 80 E_6 + \cdots , \\
(\wp(z))^3 &= \frac{1}{z^6} + \frac{9E_4}{z^2} + 15 E_6 + \cdots .
\end{aligned}$$

From these, one sees that the difference $(\wp'(z))^2 - 4(\wp(z))^3 + 60E_4\wp(z) + 140E_6$ is holomorphic near 0, and in fact equal to 0 at the origin. Since this difference is also doubly periodic, we conclude by Theorem 1.2 that it is constant, and hence identically 0. This proves the following corollary.

Corollary 2.3 *If $g_2 = 60E_4$ and $g_3 = 140E_6$, then*

$$(\wp')^2 = 4\wp^3 - g_2 \wp - g_3.$$

Note that this identity is another version of Theorem 1.7, and it allows one to express the symmetric functions of the e_j's in terms of the Eisenstein series.

2.2 Eisenstein series and divisor functions

We will describe now the link between Eisenstein series and some number-theoretic quantities. This relation comes about if we consider the Fourier coefficients in the Fourier expansion of the periodic function $E_k(\tau)$. Equivalently, we can write $\mathcal{E}(z) = E_k(\tau)$ with $z = e^{2\pi i\tau}$, and investigate the Laurent expansion of $\mathcal{E}$ as a function of z.

We begin with a lemma.

Lemma 2.4 *If* $k \geq 2$ *and* $\mathrm{Im}(\tau) > 0$, *then*

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\tau)^k} = \frac{(-2\pi i)^k}{(k-1)!} \sum_{\ell=1}^{\infty} \ell^{k-1} e^{2\pi i\tau\ell}.$$

Proof. This identity follows from applying the Poisson summation formula to $f(z) = 1/(z+\tau)^k$; see Exercise 7 in Chapter 4.

An alternate proof consists of noting that it first suffices to establish the formula for $k = 2$, since the other cases are then obtained by differentiating term by term. To prove this special case, we differentiate the formula for the cotangent derived in Chapter 5

$$\sum_{n=-\infty}^{\infty} \frac{1}{n+\tau} = \pi \cot \pi\tau.$$

This yields

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\tau)^2} = \frac{\pi^2}{\sin^2(\pi\tau)}.$$

Now use Euler's formula for the sine and the fact that

$$\sum_{r=1}^{\infty} rw^r = \frac{w}{(1-w)^2} \qquad \text{with } w = e^{2\pi i\tau}$$

to obtain the desired result.

As a consequence of this lemma, we can draw a connection between the Eisenstein series, the zeta function, and the divisor functions. The

divisor function $\sigma_\ell(r)$ that arises here is defined as the sum of the ℓth powers of the divisors of r, that is,

$$\sigma_\ell(r) = \sum_{d|r} d^\ell.$$

Theorem 2.5 *If $k \geq 4$ is even, and* $\text{Im}(\tau) > 0$, *then*

$$E_k(\tau) = 2\zeta(k) + \frac{2(-1)^{k/2}(2\pi)^k}{(k-1)!} \sum_{r=1}^{\infty} \sigma_{k-1}(r) e^{2\pi i \tau r}.$$

Proof. First observe that $\sigma_{k-1}(r) \leq r r^{k-1} = r^k$. If $\text{Im}(\tau) = t$, then whenever $t \geq t_0$ we have $|e^{2\pi i r\tau}| \leq e^{-2\pi r t_0}$, and we see that the series in the theorem is absolutely convergent in any half-plane $t \geq t_0$, by comparison with $\sum_{r=1}^{\infty} r^k e^{-2\pi r t_0}$. To establish the formula, we use the definition of E_k, that of ζ, the fact that k is even, and the previous lemma (with τ replaced by $m\tau$) to get successively

$$\begin{aligned}
E_k(\tau) &= \sum_{(n,m)\neq(0,0)} \frac{1}{(n+m\tau)^k} \\
&= \sum_{n\neq 0} \frac{1}{n^k} + \sum_{m\neq 0} \sum_{n=-\infty}^{\infty} \frac{1}{(n+m\tau)^k} \\
&= 2\zeta(k) + \sum_{m\neq 0} \sum_{n=-\infty}^{\infty} \frac{1}{(n+m\tau)^k} \\
&= 2\zeta(k) + 2\sum_{m>0} \sum_{n=-\infty}^{\infty} \frac{1}{(n+m\tau)^k} \\
&= 2\zeta(k) + 2\sum_{m>0} \frac{(-2\pi i)^k}{(k-1)!} \sum_{\ell=1}^{\infty} \ell^{k-1} e^{2\pi i m\tau\ell} \\
&= 2\zeta(k) + \frac{2(-1)^{k/2}(2\pi)^k}{(k-1)!} \sum_{m>0} \sum_{\ell=1}^{\infty} \ell^{k-1} e^{2\pi i \tau m\ell} \\
&= 2\zeta(k) + \frac{2(-1)^{k/2}(2\pi)^k}{(k-1)!} \sum_{r=1}^{\infty} \sigma_{k-1}(r) e^{2\pi i \tau r}.
\end{aligned}$$

This proves the desired formula.

Finally, we turn to the *forbidden* case $k = 2$. The series we have in mind $\sum_{(n,m)\neq(0,0)} 1/(n+m\tau)^2$ no longer converges absolutely, but we

seek to give it a meaning anyway. We define

$$F(\tau) = \sum_m \left(\sum_n \frac{1}{(n + m\tau)^2} \right)$$

summed in the indicated order with $(n, m) \neq (0, 0)$. The argument given in the above theorem proves that the double sum converges, and in fact has the expected expression.

Corollary 2.6 *The double sum defining F converges in the indicated order. We have*

$$F(\tau) = 2\zeta(2) - 8\pi^2 \sum_{r=1}^{\infty} \sigma(r) e^{2\pi i r\tau},$$

where $\sigma(r) = \sum_{d|r} d$ is the sum of the divisors of r.

It can be seen that $F(-1/\tau)\tau^{-2}$ does *not* equal $F(\tau)$, and this is the same as saying that the double series for F gives a different value ($\tilde{F}$, the reverse of F) when we sum first in m and then in n. It turns out that nevertheless the **forbidden Eisenstein series** $F(\tau)$ can be used in a crucial way in the proof of the celebrated theorem about representing an integer as the sum of four squares. We turn to these matters in the next chapter.

3 Exercises

1. Suppose that a meromorphic function f has two periods ω_1 and ω_2, with $\omega_2/\omega_1 \in \mathbb{R}$.

(a) Suppose ω_2/ω_1 is rational, say equal to p/q, where p and q are relatively prime integers. Prove that as a result the periodicity assumption is equivalent to the assumption that f is periodic with the simple period $\omega_0 = \frac{1}{q}\omega_1$. [Hint: Since p and q are relatively prime, there exist integers m and n such that $mq + np = 1$ (Corollary 1.3, Chapter 8, Book I).]

(b) If ω_2/ω_1 is irrational, then f is constant. To prove this, use the fact that $\{m - n\tau\}$ is dense in $\mathbb{R}$ whenever τ is irrational and m, n range over the integers.

2. Suppose that $a_1, \ldots, a_r$ and $b_1, \ldots, b_r$ are the zeros and poles, respectively, in the fundamental parallelogram of an elliptic function f. Show that

$$a_1 + \cdots + a_r - b_1 - \cdots - b_r = n\omega_1 + m\omega_2$$

for some integers n and m.

[Hint: If the boundary of the parallelogram contains no zeros or poles, simply integrate $zf'(z)/f(z)$ over that boundary, and observe that the integral of $f'(z)/f(z)$ over a side is an integer multiple of $2\pi i$. If there are zeros or poles on the side of the parallelogram, translate it by a small amount to reduce the problem to the first case.]

3. In contrast with the result in Lemma 1.5, prove that the series

$$\sum_{n+m\tau\in\Lambda^*} \frac{1}{|n+m\tau|^2} \qquad \text{where } \tau \in \mathbb{H}$$

does not converge. In fact, show that

$$\sum_{1\leq n^2+m^2\leq R^2} 1/(n^2+m^2) = 2\pi \log R + O(1) \qquad \text{as } R \to \infty.$$

4. By rearranging the series

$$\frac{1}{z^2} + \sum_{\omega\in\Lambda^*} \left[\frac{1}{(z+\omega)^2} - \frac{1}{\omega^2}\right],$$

show directly, without differentiation, that $\wp(z+\omega) = \wp(z)$ whenever $\omega \in \Lambda$.

[Hint: For R sufficiently large, note that $\wp(z) = \wp^R(z) + O(1/R)$, where $\wp^R(z) = z^{-2} + \sum_{0<|\omega|<R}((z+\omega)^{-2} - \omega^{-2})$. Next, observe that both $\wp^R(z+1) - \wp^R(z)$ and $\wp^R(z+\tau) - \wp^R(z)$ are $O(\sum_{R-c<|\omega|<R+c} |\omega|^{-2}) = O(1/R)$.]

5. Let $\sigma(z)$ be the canonical product

$$\sigma(z) = z\prod_{j=1}^{\infty} E_2(z/\tau_j),$$

where τ_j is an enumeration of the periods $\{n+m\tau\}$ with $(n,m) \neq (0,0)$, and $E_2(z) = (1-z)e^{z+z^2/2}$.

(a) Show that $\sigma(z)$ is an entire function of order 2 that has simple zeros at all the periods $n+m\tau$, and vanishes nowhere else.

(b) Show that

$$\frac{\sigma'(z)}{\sigma(z)} = \frac{1}{z} + \sum_{(n,m)\neq(0,0)} \left[\frac{1}{z-n-m\tau} + \frac{1}{n+m\tau} + \frac{z}{(n+m\tau)^2}\right],$$

and that this series converges whenever z is not a lattice point.

(c) Let $L(z) = -\sigma'(z)/\sigma(z)$. Then

$$L'(z) = \frac{(\sigma'(z))^2 - \sigma(z)\sigma''(z)}{(\sigma(z))^2} = \wp(z).$$

6. Prove that $\wp''$ is a quadratic polynomial in $\wp$.

7. Setting $\tau = 1/2$ in the expression

$$\sum_{m=-\infty}^{\infty} \frac{1}{(m+\tau)^2} = \frac{\pi^2}{\sin^2(\pi\tau)},$$

deduce that

$$\sum_{m\geq 1,\ m \text{ odd}} \frac{1}{m^2} = \frac{\pi^2}{8} \quad \text{and} \quad \sum_{m\geq 1} \frac{1}{m^2} = \frac{\pi^2}{6} = \zeta(2).$$

Similarly, using $\sum 1/(m+\tau)^4$ deduce that

$$\sum_{m\geq 1,\ m \text{ odd}} \frac{1}{m^4} = \frac{\pi^4}{96} \quad \text{and} \quad \sum_{m\geq 1} \frac{1}{m^4} = \frac{\pi^4}{90} = \zeta(4).$$

These results were already obtained using Fourier series in the exercises at the end of Chapters 2 and 3 in Book I.

8. Let

$$E_4(\tau) = \sum_{(n,m)\neq(0,0)} \frac{1}{(n+m\tau)^4}$$

be the Eisenstein series of order 4.

(a) Show that $E_4(\tau) \to \pi^4/45$ as $\mathrm{Im}(\tau) \to \infty$.

(b) More precisely,

$$\left| E_4(\tau) - \frac{\pi^4}{45} \right| \leq ce^{-2\pi t} \qquad \text{if } \tau = x + it \text{ and } t \geq 1.$$

(c) Deduce that

$$\left| E_4(\tau) - \tau^{-4}\frac{\pi^4}{45} \right| \leq ct^{-4}e^{-2\pi/t} \qquad \text{if } \tau = it \text{ and } 0 < t \leq 1.$$

4 Problems

1. Besides the approach in Section 1.2, there are several alternate ways of dealing with the sum $\sum 1/(z+\omega)^2$, where $\omega = n + m\tau$. For example, one may sum either (a) circularly, (b) first in n then in m, (c) or first in m then in n.

(a) Prove that if $z \notin \Lambda$, then

$$\lim_{R\to\infty} \sum_{n^2+m^2\leq R^2} \frac{1}{(z+n+m\tau)^2} = S_1(z)$$

exists and $S_1(z) = \wp(z) + c_1$.

(b) Similarly,

$$\sum_m \left(\sum_n \frac{1}{(z+n+m\tau)^2} \right) = S_2(z)$$

exists and $S_2(z) = \wp(z) + c_2$, where $c_2 = F(\tau)$, and F is the forbidden Eisenstein series.

(c) Also

$$\sum_n \left(\sum_m \frac{1}{(z+n+m\tau)^2} \right) = S_3(z)$$

exists with $S_3(z) = \wp(z) + c_3$, and $c_3 = \tilde{F}(\tau)$, the reverse of F.

[Hint: To prove (a), it suffices to show that $\lim_{R\to\infty}, \sum_{1\leq n^2+m^2\leq R^2} 1/(n+m\tau)^2 = c_1$ exists. This is proved by a comparision with $\int_{1\leq x^2+y^2\leq R^2} \frac{dx}{(x+y\tau)^2} = I(R)$. It can be shown that $I(R) = 0$, which follows because $(x+y\tau)^{-2} = -(\partial/\partial x)(x+y\tau)^{-1}$.]

2. Show that

$$\wp(z) = c + \pi^2 \sum_{m=-\infty}^{\infty} \frac{1}{\sin^2((z+m\tau)\pi)}$$

where c is an appropriate constant. In fact, by part (b) of the previous problem $c = -F(\tau)$.

3.* Suppose Ω is a simply connected domain that excludes the three roots of the polynomial $4z^3 - g_2 z - g_3$. For $\omega_0 \in \Omega$ and ω_0 fixed, define the function I on Ω by

$$I(\omega) = \int_{\omega_0}^{\omega} \frac{dz}{\sqrt{4z^3 - g_2 z - g_3}} \qquad \omega \in \Omega.$$

Then the function I has an inverse given by $\wp(z+\alpha)$ for some constant α; that is,

$$I(\wp(z+\alpha)) = z$$

for appropriate α.

[Hint: Prove that $(I(\wp(z+\alpha)))' = \pm 1$, and use the fact that $\wp$ is even.]

4.* Suppose τ is purely imaginary, say $\tau = it$ with $t > 0$. Consider the division of the complex plane into congruent rectangles obtained by considering the lines $x = n/2$, $y = tm/2$ as n and m range over the integers. (An example is the rectangle whose vertices are $0, 1/2, 1/2 + \tau/2$, and $\tau/2$.)

(a) Show that $\wp$ is real-valued on all these lines, and hence on the boundaries of all these rectangles.

(b) Prove that $\wp$ maps the interior of each rectangle conformally to the upper (or lower) half-plane.

10 Applications of Theta Functions

> The problem of the representation of an integer n as the sum of a given number k of integral squares is one of the most celebrated in the theory of numbers. Its history may be traced back to Diophantus, but begins effectively with Girard's (or Fermat's) theorem that a prime $4m + 1$ is the sum of two squares. Almost every arithmetician of note since Fermat has contributed to the solution of the problem, and it has its puzzles for us still.
>
> *G. H. Hardy,* 1940

This chapter is devoted to a closer look at the theory of theta functions and some of its applications to combinatorics and number theory.

The theta function is given by the series

$$\Theta(z|\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{2\pi i n z},$$

which converges for all $z \in \mathbb{C}$, and τ in the upper half-plane.

A remarkable feature of the theta function is its dual nature. When viewed as a function of z, we see it in the arena of elliptic functions, since Θ is periodic with period 1 and "quasi-period" τ. When considered as a function of τ, Θ reveals its modular nature and close connection with the partition function and the problem of representation of integers as sums of squares.

The two main tools allowing us to exploit these links are the triple-product for Θ and its transformation law. Once we have proved these theorems, we give a brief introduction to the connection with partitions, and then pass to proofs of the celebrated theorems about representation of integers as sums of two or four squares.

1 Product formula for the Jacobi theta function

In its most elaborate form, Jacobi's **theta function** is defined for $z \in \mathbb{C}$ and $\tau \in \mathbb{H}$ by

$$\Theta(z|\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{2\pi i n z}. \tag{1}$$

Two significant special cases (or variants) are $\theta(\tau)$ and $\vartheta(t)$, which are defined by

$$\theta(\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau}, \qquad \tau \in \mathbb{H},$$

$$\vartheta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t}, \qquad t > 0.$$

In fact, the relation between these various functions is given by $\theta(\tau) = \Theta(0|\tau)$ and $\vartheta(t) = \theta(it)$, with of course, $t > 0$.

We have already encountered these functions several times. For example, in the study of the heat diffusion equation for the circle, in Chapter 4 of Book I, we found that the heat kernel was given by

$$H_t(x) = \sum_{n=-\infty}^{\infty} e^{-4\pi^2 n^2 t} e^{2\pi i n x},$$

and therefore $H_t(x) = \Theta(x|4\pi i t)$.

Another instance was the occurence of ϑ in the study of the zeta function. In fact, we proved in Chapter 6 that the functional equation of ϑ implied that of ζ, which then led to the analytic continuation of the zeta function.

We begin our closer look at Θ as a function of z, with τ fixed, by recording its basic structural properties, which to a large extent characterize it.

Proposition 1.1 *The function Θ satisfies the following properties:*

(i) Θ *is entire in $z \in \mathbb{C}$ and holomorphic in $\tau \in \mathbb{H}$.*

(ii) $\Theta(z+1|\tau) = \Theta(z|\tau)$.

(iii) $\Theta(z+\tau|\tau) = \Theta(z|\tau)e^{-\pi i \tau}e^{-2\pi i z}$.

(iv) $\Theta(z|\tau) = 0$ *whenever $z = 1/2 + \tau/2 + n + m\tau$ and $n, m \in \mathbb{Z}$.*

Proof. Suppose that $\text{Im}(\tau) = t \geq t_0 > 0$ and $z = x + iy$ belongs to a bounded set in $\mathbb{C}$, say $|z| \leq M$. Then, the series defining Θ is absolutely and uniformly convergent, since

$$\sum_{n=-\infty}^{\infty} |e^{\pi i n^2 \tau} e^{2\pi i n z}| \leq C \sum_{n \geq 0} e^{-\pi n^2 t_0} e^{2\pi n M} < \infty.$$

Therefore, for each fixed $\tau \in \mathbb{H}$ the function $\Theta(\cdot|\tau)$ is entire, and for each fixed $z \in \mathbb{C}$ the function $\Theta(z|\cdot)$ is holomorphic in the upper half-plane.

Since the exponential $e^{2\pi i n z}$ is periodic of period 1, property (ii) is immediate from the definition of Θ.

To show the third property we may complete the squares in the expression for $\Theta(z + \tau|\tau)$. In detail, we have

$$\begin{aligned}
\Theta(z+\tau|\tau) &= \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{2\pi i n (z+\tau)} \\
&= \sum_{n=-\infty}^{\infty} e^{\pi i (n^2+2n)\tau} e^{2\pi i n z} \\
&= \sum_{n=-\infty}^{\infty} e^{\pi i (n+1)^2 \tau} e^{-\pi i \tau} e^{2\pi i n z} \\
&= \sum_{n=-\infty}^{\infty} e^{\pi i (n+1)^2 \tau} e^{-\pi i \tau} e^{2\pi i (n+1) z} e^{-2\pi i z} \\
&= \Theta(z|\tau) e^{-\pi i \tau} e^{-2\pi i z}.
\end{aligned}$$

Thus we see that $\Theta(z|\tau)$, as a function of z, is periodic with period 1 and "quasi-periodic" with period τ.

To establish the last property it suffices, by what was just shown, to prove that $\Theta(1/2 + \tau/2|\tau) = 0$. Again, we use the interplay between n and n^2 to get

$$\begin{aligned}
\Theta(1/2+\tau/2|\tau) &= \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{2\pi i n (1/2+\tau/2)} \\
&= \sum_{n=-\infty}^{\infty} (-1)^n e^{\pi i (n^2+n)\tau}.
\end{aligned}$$

To see that this last sum is identically zero, it suffices to match $n \geq 0$ with $-n-1$, and to observe that they have opposite parity, and that $(-n-1)^2 + (-n-1) = n^2 + n$. This completes the proof of the proposition.

We consider next a product $\Pi(z|\tau)$ that enjoys the same structural properties as $\Theta(z|\tau)$ as a function of z. This product is defined for $z \in \mathbb{C}$ and $\tau \in \mathbb{H}$ by

$$\Pi(z|\tau) = \prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1}e^{2\pi iz})(1+q^{2n-1}e^{-2\pi iz}),$$

where we have used the notation that is standard in the subject, namely $q = e^{\pi i\tau}$. The function $\Pi(z|\tau)$ is sometimes referred to as the **triple-product**.

Proposition 1.2 *The function* $\Pi(z|\tau)$ *satisfies the following properties:*

(i) $\Pi(z,\tau)$ *is entire in* $z \in \mathbb{C}$ *and holomorphic for* $\tau \in \mathbb{H}$.

(ii) $\Pi(z+1|\tau) = \Pi(z|\tau)$.

(iii) $\Pi(z+\tau|\tau) = \Pi(z|\tau)e^{-\pi i\tau}e^{-2\pi iz}$.

(iv) $\Pi(z|\tau) = 0$ *whenever* $z = 1/2 + \tau/2 + n + m\tau$ *and* $n, m \in \mathbb{Z}$. *Moreover, these points are simple zeros of* $\Pi(\cdot|\tau)$, *and* $\Pi(\cdot|\tau)$ *has no other zeros.*

Proof. If $\mathrm{Im}(\tau) = t \geq t_0 > 0$ and $z = x+iy$, then $|q| \leq e^{-\pi t_0} < 1$ and

$$(1-q^{2n})(1+q^{2n-1}e^{2\pi iz})(1+q^{2n-1}e^{-2\pi iz}) = 1 + O\left(|q|^{2n-1}e^{2\pi|z|}\right).$$

Since the series $\sum |q|^{2n-1}$ converges, the results for infinite products in Chapter 5 guarantee that $\Pi(z|\tau)$ defines an entire function of z with $\tau \in \mathbb{H}$ fixed, and a holomorphic function for $\tau \in \mathbb{H}$ with $z \in \mathbb{C}$ fixed.

Also, it is clear from the definition that $\Pi(z|\tau)$ is periodic of period 1 in the z variable.

To prove the third property, we first observe that since $q^2 = e^{2\pi i\tau}$ we have

$$\begin{aligned}\Pi(z+\tau|\tau) &= \prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1}e^{2\pi i(z+\tau)})(1+q^{2n-1}e^{-2\pi i(z+\tau)}) \\ &= \prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n+1}e^{2\pi iz})(1+q^{2n-3}e^{-2\pi iz}).\end{aligned}$$

Comparing this last product with $\Pi(z|\tau)$, and isolating the factors that are either missing or extra leads to

$$\Pi(z+\tau|\tau) = \Pi(z|\tau)\left(\frac{1+q^{-1}e^{-2\pi iz}}{1+qe^{2\pi iz}}\right).$$

Hence (iii) follows because $(1+x)/(1+x^{-1}) = x$, whenever $x \neq -1$.

Finally, to find the zeros of $\Pi(z|\tau)$ we recall that a product that converges vanishes only if at least one of its factors is zero. Clearly, the factor $(1-q^n)$ never vanishes since $|q|<1$. The second factor $(1+q^{2n-1}e^{2\pi iz})$ vanishes when $q^{2n-1}e^{2\pi iz} = -1 = e^{\pi i}$. Since $q = e^{\pi i\tau}$, we then have[1]

$$(2n-1)\tau + 2z = 1 \pmod{2}.$$

Hence,

$$z = 1/2 + \tau/2 - n\tau \pmod{1},$$

and this takes care of the zeros of the type $1/2+\tau/2-n\tau+m$ with $n \geq 1$ and $m \in \mathbb{Z}$. Similarly, the third factor vanishes if

$$(2n-1)\tau - 2z = 1 \pmod{2}$$

which implies that

$$\begin{aligned} z &= -1/2 - \tau/2 + n\tau \pmod{1} \\ &= 1/2 + \tau/2 + n'\tau \pmod{1}, \end{aligned}$$

where $n' \geq 0$. This exhausts the zeros of $\Pi(\cdot|\tau)$. Finally, these zeros are simple, since the function $e^w - 1$ vanishes at the origin to order 1 (a fact obvious from a power series expansion or a simple differentiation).

The importance of the product Π comes from the following theorem, called the product formula for the theta function. The fact that $\Theta(z|\tau)$ and $\Pi(z|\tau)$ satisfy similar properties hints at a close connection between the two. This is indeed the case.

Theorem 1.3 (Product formula) *For all $z \in \mathbb{C}$ and $\tau \in \mathbb{H}$ we have the identity $\Theta(z|\tau) = \Pi(z|\tau)$.*

Proof. Fix $\tau \in \mathbb{H}$. We claim first that there exists a constant $c(\tau)$ such that

$$\Theta(z|\tau) = c(\tau)\Pi(z|\tau). \tag{2}$$

In fact, consider the quotient $F(z) = \Theta(z|\tau)/\Pi(z|\tau)$, and note that by the previous two propositions, the function F is entire and doubly periodic with periods 1 and τ. This implies that F is constant as claimed.

[1] We use the standard short-hand, $a = b \pmod{c}$, to mean that $a-b$ is an integral multiple of c.

We must now prove that $c(\tau) = 1$ for all τ, and the main point is to establish that $c(\tau) = c(4\tau)$. If we put $z = 1/2$ in (2), so that $e^{2i\pi z} = e^{-2i\pi z} = -1$, we obtain

$$\begin{aligned}\sum_{n=-\infty}^{\infty} (-1)^n q^{n^2} &= c(\tau) \prod_{n=1}^{\infty} (1 - q^{2n})(1 - q^{2n-1})(1 - q^{2n-1}) \\ &= c(\tau) \prod_{n=1}^{\infty} \left[(1 - q^{2n-1})(1 - q^{2n})\right] (1 - q^{2n-1}) \\ &= c(\tau) \prod_{n=1}^{\infty} (1 - q^n)(1 - q^{2n-1}).\end{aligned}$$

Hence

$$c(\tau) = \frac{\sum_{n=-\infty}^{\infty} (-1)^n q^{n^2}}{\prod_{n=1}^{\infty} (1 - q^n)(1 - q^{2n-1})}. \tag{3}$$

Next, we put $z = 1/4$ in (2), so that $e^{2i\pi z} = i$. On the one hand, we have

$$\Theta(1/4|\tau) = \sum_{n=-\infty}^{\infty} q^{n^2} i^n,$$

and due to the fact that $1/i = -i$, only the terms corresponding to $n =$ even $= 2m$ are not cancelled; thus

$$\Theta(1/4|\tau) = \sum_{m=-\infty}^{\infty} q^{4m^2} (-1)^m.$$

On the other hand,

$$\begin{aligned}\Pi(1/4|\tau) &= \prod_{m=1}^{\infty} (1 - q^{2m})(1 + iq^{2m-1})(1 - iq^{2m-1}) \\ &= \prod_{m=1}^{\infty} (1 - q^{2m})(1 + q^{4m-2}) \\ &= \prod_{n=1}^{\infty} (1 - q^{4n})(1 - q^{8n-4}),\end{aligned}$$

where the last line is obtained by considering separately the two cases $2m = 4n - 4$ and $2m = 4n - 2$ in the first factor. Hence

$$c(\tau) = \frac{\sum_{n=-\infty}^{\infty} (-1)^n q^{4n^2}}{\prod_{n=1}^{\infty} (1 - q^{4n})(1 - q^{8n-4})}, \tag{4}$$

and combining (3) and (4) establishes our claim that $c(\tau) = c(4\tau)$. Successive applications of this identity give $c(\tau) = c(4^k\tau)$, and since $q^{4^k} = e^{i\pi 4^k \tau} \to 0$ as $k \to \infty$, we conclude from (2) that $c(\tau) = 1$. This proves the theorem.

The product formula for the function Θ specializes to its variant $\theta(\tau) = \Theta(0|\tau)$, and this provides a proof that θ is non-vanishing in the upper half-plane.

Corollary 1.4 *If* $\text{Im}(\tau) > 0$ *and* $q = e^{\pi i \tau}$, *then*

$$\theta(\tau) = \prod_{n=1}^{\infty}(1 - q^{2n})(1 + q^{2n-1})^2.$$

Thus $\theta(\tau) \neq 0$ *for* $\tau \in \mathbb{H}$.

The next corollary shows that the properties of the function Θ now yield the construction of an elliptic function (which is in fact closely related to the Weierstrass $\wp$ function).

Corollary 1.5 *For each fixed* $\tau \in \mathbb{H}$, *the quotient*

$$(\log \Theta(z|\tau))'' = \frac{\Theta(z|\tau)\Theta''(z|\tau) - (\Theta'(z|\tau))^2}{\Theta(z|\tau)^2}$$

is an elliptic function of order 2 *with periods* 1 *and* τ, *and with a double pole at* $z = 1/2 + \tau/2$.

In the above, the primes $'$ denote differentiation with respect to the z variable.

Proof. Let $F(z) = (\log \Theta(z|\tau))' = \Theta(z|\tau)'/\Theta(z|\tau)$. Differentiating the identities (ii) and (iii) of Proposition 1.1 gives $F(z+1) = F(z)$, $F(z+\tau) = F(z) - 2\pi i$, and differentiating again shows that $F'(z)$ is doubly periodic. Since $\Theta(z|\tau)$ vanishes only at $z = 1/2 + \tau/2$ in the fundamental parallelogram, the function $F(z)$ has only a single pole, and thus $F'(z)$ has only a double pole there.

The precise connection between $(\log \Theta(z|\tau))''$ and $\wp_\tau(z)$ is stated in Exercise 1.

For an analogy between Θ and the Weierstrass σ function, see Exercise 5 of the previous chapter.

1.1 Further transformation laws

We now come to the study of the transformation relations in the τ-variable, that is, to the modular character of Θ.

Recall that in the previous chapter, the modular character of the Weierstrass $\wp$ function and Eisenstein series E_k was reflected by the two transformations

$$\tau \mapsto \tau + 1 \quad \text{and} \quad \tau \mapsto -1/\tau,$$

which preserve the upper half-plane. In what follows, we shall denote these two transformations by T_1 and S, respectively.

When looking at the Θ function, however, it will be natural to consider instead the transformations

$$T_2 : \tau \mapsto \tau + 2 \quad \text{and} \quad S : \tau \mapsto -1/\tau,$$

since $\Theta(z|\tau + 2) = \Theta(z|\tau)$, but $\Theta(z|\tau + 1) \neq \Theta(z|\tau)$.

Our first task is to study the transformation of $\Theta(z|\tau)$ under the mapping $\tau \mapsto -1/\tau$.

Theorem 1.6 *If $\tau \in \mathbb{H}$, then*

$$\Theta(z| - 1/\tau) = \sqrt{\frac{\tau}{i}} e^{\pi i \tau z^2} \Theta(z\tau|\tau) \quad \textit{for all } z \in \mathbb{C}. \tag{5}$$

Here $\sqrt{\tau/i}$ denotes the branch of the square root defined on the upper half-plane, that is positive when $\tau = it$, $t > 0$.

Proof. It suffices to prove this formula for $z = x$ real and $\tau = it$ with $t > 0$, since for each fixed $x \in \mathbb{R}$, the two sides of equation (5) are holomorphic functions in the upper half-plane which then agree on the positive imaginary axis, and hence must be equal everywhere. Also, for a fixed $\tau \in \mathbb{H}$ the two sides define holomorphic functions in z that agree on the real axis, and hence must be equal everywhere.

With x real and $\tau = it$ the formula becomes

$$\sum_{n=-\infty}^{\infty} e^{-\pi n^2/t} e^{2\pi i n x} = t^{1/2} e^{-\pi t x^2} \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t} e^{-2\pi n x t}.$$

Replacing x by a, we find that we must prove

$$\sum_{n=-\infty}^{\infty} e^{-\pi t(n+a)^2} = \sum_{n=-\infty}^{\infty} t^{-1/2} e^{-\pi n^2/t} e^{2\pi i n a}.$$

However, this is precisely equation (3) in Chapter 4, which was derived from the Poisson summation formula.

In particular, by setting $z = 0$ in the theorem, we find the following.

Corollary 1.7 *If* $\mathrm{Im}(\tau) > 0$*, then* $\theta(-1/\tau) = \sqrt{\tau/i}\ \theta(\tau)$*.*

Note that if $\tau = it$, then $\theta(\tau) = \vartheta(t)$, and the above relation is precisely the functional equation for ϑ which appeared in Chapter 4.

The transformation law $\theta(-1/\tau) = (\tau/i)^{1/2}\theta(\tau)$ gives us very precise information about the behavior when $\tau \to 0$. The next corollary will be used later, when we need to analyze the behavior of $\theta(\tau)$ as $\tau \to 1$.

Corollary 1.8 *If* $\tau \in \mathbb{H}$*, then*

$$\begin{aligned}\theta(1 - 1/\tau) &= \sqrt{\frac{\tau}{i}} \sum_{n=-\infty}^{\infty} e^{\pi i (n+1/2)^2 \tau} \\ &= \sqrt{\frac{\tau}{i}} \left(2e^{\pi i \tau/4} + \cdots\right).\end{aligned}$$

The second identity means that $\theta(1 - 1/\tau) \sim \sqrt{\tau/i} 2e^{i\pi\tau/4}$ *as* $\mathrm{Im}(\tau) \to \infty$*.*

Proof. First, we note that n and n^2 have the same parity, so

$$\theta(1 + \tau) = \sum_{n=-\infty}^{\infty} (-1)^n e^{i\pi n^2 \tau} = \Theta(1/2|\tau),$$

hence $\theta(1 - 1/\tau) = \Theta(1/2| - 1/\tau)$. Next, we use Theorem 1.6 with $z = 1/2$, and the result is

$$\begin{aligned}\theta(1 - 1/\tau) &= \sqrt{\frac{\tau}{i}} e^{\pi i \tau/4} \Theta(\tau/2|\tau) \\ &= \sqrt{\frac{\tau}{i}} e^{\pi i \tau/4} \sum_{n=-\infty}^{\infty} e^{\pi i n^2 \tau} e^{\pi i n \tau} \\ &= \sqrt{\frac{\tau}{i}} \sum_{n=-\infty}^{\infty} e^{\pi i (n+1/2)^2 \tau}.\end{aligned}$$

The terms corresponding to $n = 0$ and $n = -1$ contribute $2e^{\pi i \tau/4}$, which has absolute value $2e^{-\pi t/4}$ where $\tau = \sigma + it$. Finally, the sum of the other terms $n \neq 0, -1$ is of order

$$O\left(\sum_{k=1}^{\infty} e^{-(k+1/2)^2 \pi t}\right) = O\left(e^{-9\pi t/4}\right).$$

Our final corollary of the transformation law pertains to the **Dedekind eta function**, which is defined for $\mathrm{Im}(\tau) > 0$ by

$$\eta(\tau) = e^{\frac{\pi i \tau}{12}} \prod_{n=1}^{\infty} (1 - e^{2\pi i n \tau}).$$

The functional equation for η given below will be relevant to our discussion of the four-square theorem, and in the theory of partitions.

Proposition 1.9 *If* $\mathrm{Im}(\tau) > 0$, *then* $\eta(-1/\tau) = \sqrt{\tau/i}\, \eta(\tau)$.

This identity is deduced by differentiating the relation in Theorem 1.6 and evaluating it at $z_0 = 1/2 + \tau/2$. The details are as follows.

Proof. From the product formula for the theta function, we may write with $q = e^{\pi i \tau}$,

$$\Theta(z|\tau) = (1 + qe^{-2\pi i z}) \prod_{n=1}^{\infty} (1 - q^{2n})(1 + q^{2n-1}e^{2\pi i z})(1 + q^{2n+1}e^{-2\pi i z}),$$

and since the first factor vanishes at $z_0 = 1/2 + \tau/2$, we see that

$$\Theta'(z_0|\tau) = 2\pi i H(\tau), \qquad \text{where } H(\tau) = \textstyle\prod_{n=1}^{\infty}(1 - e^{2\pi i n \tau})^3.$$

Next, we observe that with $-1/\tau$ replaced by τ in (5), we obtain

$$\Theta(z|\tau) = \sqrt{i/\tau} e^{-\pi i z^2/\tau} \Theta(-z/\tau| - 1/\tau).$$

If we differentiate this expression and then evaluate it at the point $z_0 = 1/2 + \tau/2$, we find

$$2\pi i H(\tau) = \sqrt{i/\tau} e^{-\frac{\pi i}{4\tau}} e^{-\frac{\pi i}{2}} e^{-\frac{\pi i \tau}{4}} \left(\frac{-2\pi i}{\tau}\right) H(-1/\tau).$$

Hence

$$e^{\frac{\pi i \tau}{4}} H(\tau) = \left(\frac{i}{\tau}\right)^{3/2} e^{-\frac{\pi i}{4\tau}} H(-1/\tau).$$

We note that when $\tau = it$, with $t > 0$, the function $\eta(\tau)$ is positive, and thus taking the cube root of the above gives $\eta(\tau) = \sqrt{i/\tau}\, \eta(-1/\tau)$; therefore this identity holds for all $\tau \in \mathbb{H}$ by analytic continuation.

A connection between the function η and the theory of elliptic functions is given in Problem 5.

2 Generating functions

Given a sequence $\{F_n\}_{n=0}^{\infty}$, which may arise either combinatorially, recursively, or in terms of some number-theoretic law, an important tool in its study is the passage to its **generating function**, defined by

$$F(x) = \sum_{n=0}^{\infty} F_n x^n.$$

Often times, the defining properties of the sequence $\{F_n\}$ imply interesting algebraic or analytic properties of the function $F(x)$, and exploiting these can eventually lead us back to new insights about the sequence $\{F_n\}$. A very simple-minded example is given by the Fibonacci sequence. (See Exercise 2). Here we want to study less elementary examples of this idea, related to the Θ function.

We shall first discuss very briefly the theory of partitions.

The **partition function** is defined as follows: if n is a positive integer, we let $p(n)$ denote the numbers of ways n can be written as a sum of positive integers. For instance, $p(1) = 1$, and $p(2) = 2$ since $2 = 2 + 0 = 1 + 1$. Also, $p(3) = 3$ since $3 = 3 + 0 = 2 + 1 = 1 + 1 + 1$. We set $p(0) = 1$ and collect some further values of $p(n)$ in the following table.

n	0	1	2	3	4	5	6	7	8	$\cdots$	12
$p(n)$	1	1	2	3	5	7	11	15	22	$\cdots$	77

The first theorem is Euler's identity for the generating function of the partition sequence $\{p(n)\}$, which is reminiscent of the product formula for the zeta function.

Theorem 2.1 *If* $|x| < 1$, *then* $\displaystyle\sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty} \frac{1}{1 - x^k}$.

Formally, we can write each fraction as

$$\frac{1}{1 - x^k} = \sum_{m=0}^{\infty} x^{km},$$

and multiply these out together to obtain $p(n)$ as the coefficient of x^n. Indeed, when we group together equal integers in a partition of n, this partition can be written as

$$n = m_1 k_1 + \cdots + m_r k_r,$$

where $k_1, \ldots, k_r$ are distinct positive integers. This partition corresponds to the term

$$(x^{k_1})^{m_1} \cdots (x^{k_r})^{m_r}$$

that arises in the product.

The justification of this formal argument proceeds as in the proof of the product formula for the zeta function (Section 1, Chapter 7); this is based on the convergence of the product $\prod 1/(1-x^k)$. This convergence in turn follows from the fact that for each fixed $|x| < 1$ one has

$$\frac{1}{1-x^k} = 1 + O(x^k).$$

A similar argument shows that the product $\prod 1/(1-x^{2n-1})$ is equal to the generating function for $p_o(n)$, the number of partitions of n into odd parts. Also, $\prod(1+x^n)$ is the generating function for $p_u(n)$, the number of partitions of n into unequal parts. Remarkably, $p_o(n) = p_u(n)$ for all n, and this translates into the identity

$$\prod_{n=1}^{\infty} \left(\frac{1}{1-x^{2n-1}} \right) = \prod_{n=1}^{\infty} (1+x^n).$$

To prove this note that $(1+x^n)(1-x^n) = 1-x^{2n}$, and therefore

$$\prod_{n=1}^{\infty}(1+x^n) \prod_{n=1}^{\infty}(1-x^n) = \prod_{n=1}^{\infty}(1-x^{2n}).$$

Moreover, taking into account the parity of integers, we have

$$\prod_{n=1}^{\infty}(1-x^{2n}) \prod_{n=1}^{\infty}(1-x^{2n-1}) = \prod_{n=1}^{\infty}(1-x^n),$$

which combined with the above proves the desired identity.

The proposition that follows is deeper, and in fact involves the Θ function directly. Let $p_{e,u}(n)$ denote the number of partitions of n into an even number of unequal parts, and $p_{o,u}(n)$ the number of partitions of n into an odd number of unequal parts. Then, Euler proved that, unless n is a pentagonal number, one has $p_{e,u}(n) = p_{o,u}(n)$. By definition, **pentagonal numbers**[2] are integers n of the form $k(3k+1)/2$, with $k \in \mathbb{Z}$. For

[2]The traditional definition is as follows. Integers of the form $n = k(k-1)/2$, $k \in \mathbb{Z}$, are "triangular numbers"; those of the form $n = k^2$ are "squares"; and those of the form $k(3k+1)/2$ are "pentagonal numbers." In general, numbers of the form $(k/2)((\ell-2)k + \ell - 4)$ are associated with an ℓ-sided polygon.

example, the first few pentagonal numbers are $1, 2, 5, 7, 12, 15, 22, 26, \ldots$. In fact, if n is pentagonal, then

$$p_{e,u}(n) - p_{o,u}(n) = (-1)^k, \qquad \text{if } n = k(3k+1)/2.$$

To prove this result, we first observe that

$$\prod_{n=1}^{\infty}(1 - x^n) = \sum_{n=1}^{\infty}[p_{e,u}(n) - p_{o,u}(n)]x^n.$$

This follows since multiplying the terms in the product, we obtain terms of the form $(-1)^r x^{n_1+\cdots+n_r}$ where the integers $n_1, \ldots, n_r$ are distinct. Hence in the coefficient of x^n, each partition $n_1 + \cdots + n_r$ of n into an even number of unequal parts contributes for $+1$ (r is even), and each partition into an odd number of unequal parts contributes -1 (r is odd). This gives precisely the coefficient $p_{e,u}(n) - p_{o,u}(n)$.

With the above identity, we see that Euler's theorem is a consequence of the following proposition.

Proposition 2.2 $\displaystyle\prod_{n=1}^{\infty}(1 - x^n) = \sum_{k=-\infty}^{\infty}(-1)^k x^{\frac{k(3k+1)}{2}}.$

Proof. If we set $x = e^{2\pi i u}$, then we can write

$$\prod_{n=1}^{\infty}(1 - x^n) = \prod_{n=1}^{\infty}(1 - e^{2\pi i n u})$$

in terms of the triple product

$$\prod_{n=1}^{\infty}(1 - q^{2n})(1 + q^{2n-1}e^{2\pi i z})(1 + q^{2n-1}e^{-2\pi i z})$$

by letting $q = e^{3\pi i u}$ and $z = 1/2 + u/2$. This is because

$$\prod_{n=1}^{\infty}(1 - e^{2\pi i 3nu})(1 - e^{2\pi i(3n-1)u})(1 - e^{2\pi i(3n-2)u}) = \prod_{n=1}^{\infty}(1 - e^{2\pi i n u}).$$

By Theorem 1.3 the product equals

$$\sum_{n=-\infty}^{\infty} e^{3\pi i n^2 u}(-1)^n e^{2\pi i n u/2} = \sum_{n=-\infty}^{\infty}(-1)^n e^{\pi i n(3n+1)u}$$

$$= \sum_{n=-\infty}^{\infty} (-1)^n x^{n(3n+1)/2},$$

which was to be proved.

We make a final comment about the partition function $p(n)$. The nature of its growth as $n \to \infty$ can be analyzed in terms of the behavior of $1/\prod_{n=1}^{\infty}(1-x)^n$ as $|x| \to 1$. In fact, by elementary considerations, we can get a rough order of growth of $p(n)$ from the growth of the generating function as $x \to 1$; see Exercises 5 and 6. A more refined analysis requires the transformation properties of the generating function which goes back to the corresponding Proposition 1.9 about η. This leads to a very good asymptotic formula for $p(n)$. It may be found in Appendix A.

3 The theorems about sums of squares

The ancient Greeks were fascinated by triples of integers (a, b, c) that occurred as sides of right triangles. These are the "Pythagorean triples," which satisfy $a^2 + b^2 = c^2$. According to Diophantus of Alexandria (ca. 250 AD), if c is an integer of the above kind, and a and b have no common factors (a case to which one may easily reduce), then c is the sum of two squares, that is, $c = m^2 + n^2$ with $m, n \in \mathbb{Z}$; and conversely, any such c arises as the hypotenuse of a triangle whose sides are given by a Pythagorean triple (a, b, c). (See Exercise 8.) Therefore, it is natural to ask the following question: which integers can be written as the sum of two squares? It is easy to see that no number of the form $4k + 3$ can be so written, but to determine which integers can be expressed in this way is not obvious.

Let us pose the question in a more quantitative form. We define $r_2(n)$ to be the number of ways n can be written as the sum of two squares, counting obvious repetitions; that is, $r_2(n)$ is the number of pairs (x, y), $x, y \in \mathbb{Z}$, so that

$$n = x^2 + y^2.$$

For example, $r_2(3) = 0$, but $r_2(5) = 8$ because $5 = (\pm 2)^2 + (\pm 1)^2$, and also $5 = (\pm 1)^2 + (\pm 2)^2$. Hence, our first problem can be posed as follows:

> **Sum of two squares:** Which integers can be written as a sum of two squares? More precisely, can one determine an expression for $r_2(n)$?

Next, since not every positive integer can be expressed as the sum of two squares, we may ask if three squares, or possibly four squares suffice.

However, the fact is that there are infinitely many integers that cannot be written as the sum of three squares, since it is easy to check that no integer of the form $8k + 7$ can be so written. So we turn to the question of four squares and define, in analogy with $r_2(n)$, the function $r_4(n)$ to be the number of ways of expressing n as a sum of four squares. Therefore, a second problem that arises is:

> **Sum of four squares:** Can every positive integer be written as a sum of four squares? More precisely, determine a formula for $r_4(n)$.

It turns out that the problems of two squares and four squares, which go back to the third century, were not resolved until about 1500 years later, and their full solution was first given by the use of Jacobi's theory of theta functions!

3.1 The two-squares theorem

The problem of representing an integer as the sum of two squares, while obviously additive in nature, has a nice multiplicative aspect: if n and m are two integers that can be written as the sum of two squares, then so can their product nm. Indeed, suppose $n = a^2 + b^2$, $m = c^2 + d^2$, and consider the complex number

$$x + iy = (a + ib)(c + id).$$

Clearly, x and y are integers since $a, b, c, d \in \mathbb{Z}$, and by taking absolute values on both sides we see that

$$x^2 + y^2 = (a^2 + b^2)(c^2 + d^2),$$

and it follows that $nm = x^2 + y^2$.

For these reasons the divisibility properties of n play a crucial role in determining $r_2(n)$. To state the basic result we define two new **divisor functions**: we let $d_1(n)$ denote the number of divisors of n of the form $4k + 1$, and $d_3(n)$ the number of divisors of n of the form $4k + 3$. The main result of this section provides a complete answer to the two-squares problem:

Theorem 3.1 *If $n \geq 1$, then $r_2(n) = 4(d_1(n) - d_3(n))$.*

A direct consequence of the above formula for $r_2(n)$ may be stated as follows. If $n = p_1^{a_1} \cdots p_r^{a_r}$ is the prime factorization of n where $p_1, \ldots, p_r$ are distinct, then:

The positive integer n can be represented as the sum of two squares if and only if every prime p_j of the form $4k+3$ that occurs in the factorization of n has an even exponent a_j.

The proof of this deduction is outlined in Exercise 9.

To prove the theorem, we first establish a crucial relationship that identifies the generating function of the sequence $\{r_2(n)\}_{n=1}^{\infty}$ with the square of the θ function, namely

$$\theta(\tau)^2 = \sum_{n=0}^{\infty} r_2(n) q^n, \tag{6}$$

whenever $q = e^{\pi i \tau}$ with $\tau \in \mathbb{H}$. The proof of this identity relies simply on the definition of r_2 and θ. Indeed, if we first recall that $\theta(\tau) = \sum_{-\infty}^{\infty} q^{n^2}$, then we obtain

$$\begin{aligned}
\theta(\tau)^2 &= \left(\sum_{n_1=-\infty}^{\infty} q^{n_1^2}\right)\left(\sum_{n_2=-\infty}^{\infty} q^{n_2^2}\right) \\
&= \sum_{(n_1,n_2)\in\mathbb{Z}\times\mathbb{Z}} q^{n_1^2+n_2^2} \\
&= \sum_{n=0}^{\infty} r_2(n) q^n,
\end{aligned}$$

since $r_2(n)$ counts the number of pairs (n_1, n_2) with $n_1^2 + n_2^2 = n$.

Proposition 3.2 *The identity $r_2(n) = 4(d_1(n) - d_3(n))$, $n \geq 1$, is equivalent to the identities*

$$\theta(\tau)^2 = 2\sum_{n=-\infty}^{\infty} \frac{1}{q^n + q^{-n}} = 1 + 4\sum_{n=1}^{\infty} \frac{q^n}{1+q^{2n}}, \tag{7}$$

whenever $q = e^{\pi i \tau}$ and $\tau \in \mathbb{H}$.

Proof. We note first that both series converge absolutely since $|q| < 1$, and the first equals the second, because $1/(q^n + q^{-n}) = q^{|n|}/(1 + q^{2|n|})$.

Since $(1+q^{2n})^{-1} = (1-q^{2n})/(1-q^{4n})$, the right-hand side of (7) equals

$$1 + 4\sum_{n=1}^{\infty}\left(\frac{q^n}{1-q^{4n}} - \frac{q^{3n}}{1-q^{4n}}\right).$$

However, since $1/(1-q^{4n}) = \sum_{m=0}^{\infty} q^{4nm}$, we have

$$\sum_{n=1}^{\infty} \frac{q^n}{1-q^{4n}} = \sum_{n=1}^{\infty}\sum_{m=0}^{\infty} q^{n(4m+1)} = \sum_{k=1}^{\infty} d_1(k)q^k,$$

because $d_1(k)$ counts the number of divisors of k that are of the form $4m+1$. Observe that the series $\sum d_1(k)q^k$ converges since $d_1(k) \leq k$.

A similar argument shows that

$$\sum_{n=1}^{\infty} \frac{q^{3n}}{1-q^{4n}} = \sum_{k=1}^{\infty} d_3(k)q^k,$$

and the proof of the proposition is complete.

In effect, we see that the identity (6) links the original problem in arithmetic with the problem in complex analysis of establishing the relation (7).

We shall now find it convenient to use $\mathcal{C}(\tau)$ to denote[3]

$$\mathcal{C}(\tau) = 2\sum_{n=-\infty}^{\infty} \frac{1}{q^n + q^{-n}} = \sum_{n=-\infty}^{\infty} \frac{1}{\cos(n\pi\tau)}, \tag{8}$$

where $q = e^{\pi i \tau}$ and $\tau \in \mathbb{H}$. Our work then becomes to prove the identity $\theta(\tau)^2 = \mathcal{C}(\tau)$.

What is truly remarkable are the different yet parallel ways that the functions θ and $\mathcal{C}$ arise. The genesis of the function θ may be thought to be the heat diffusion equation on the real line; the corresponding heat kernel is given in terms of the Gaussian $e^{-\pi x^2}$ which is its own Fourier transform; and finally the transformation rule for θ results from the Poisson summation formula.

The parallel with $\mathcal{C}$ is that it arises from another differential equation: the steady-state heat equation in a strip; there, the corresponding kernel is $1/\cosh \pi x$ (Section 1.3, Chapter 8), which again is its own Fourier transform (Example 3, Chapter 3). The transformation rule for $\mathcal{C}$ results, once again, from the Poisson summation formula.

To prove the identity $\theta^2 = \mathcal{C}$ we will first show that these two functions satisfy the same structural properties. For θ^2 we had the transformation law $\theta(\tau)^2 = (i/\tau)\theta(-1/\tau)^2$ (Corollary 1.7).

---

[3]We denote the function by $\mathcal{C}$ because we are summing a series of cosines.

An identical transformation law holds for $\mathcal{C}(\tau)$! Indeed, if we set $a = 0$ in the relation (5) of Chapter 4 we obtain

$$\sum_{n=-\infty}^{\infty} \frac{1}{\cosh(\pi n t)} = \frac{1}{t} \sum_{n=-\infty}^{\infty} \frac{1}{\cosh(\pi n/t)}.$$

This is precisely the identity

$$\mathcal{C}(\tau) = (i/\tau)\,\mathcal{C}(-1/\tau)$$

for $\tau = it$, $t > 0$, which therefore also holds for all $\tau \in \mathbb{H}$ by analytic continuation.

It is also obvious from their definitions that both $\theta(\tau)^2$ and $\mathcal{C}(\tau)$ tend to 1 as $\mathrm{Im}(\tau) \to \infty$. The last property we want to examine is the behavior of both functions at the "cusp" $\tau = 1$.[4]

For θ^2 we shall invoke Corollary 1.8 to see that $\theta(1 - 1/\tau)^2 \sim 4(\tau/i)e^{\pi i \tau/2}$ as $\mathrm{Im}(\tau) \to \infty$.

For $\mathcal{C}$ we can do the same, again using the Poisson summation formula. In fact, if we set $a = 1/2$ in equation (5), Chapter 4, we find

$$\sum_{n=-\infty}^{\infty} \frac{(-1)^n}{\cosh(\pi n/t)} = t \sum_{n=-\infty}^{\infty} \frac{1}{\cosh(\pi(n+1/2)t)}.$$

Therefore, by analytic continuation we deduce that

$$\mathcal{C}(1 - 1/\tau) = \left(\frac{\tau}{i}\right) \sum_{n=-\infty}^{\infty} \frac{1}{\cos(\pi(n+1/2)\tau)}.$$

The main terms of this sum are those for $n = -1$ and $n = 0$. This easily gives

$$\mathcal{C}(1 - 1/\tau) = 4\left(\frac{\tau}{i}\right) e^{\pi i \tau/2} + O\left(|\tau| e^{-3\pi t/2}\right), \qquad \text{as } t \to \infty,$$

and where $\tau = \sigma + it$. We summarize our conclusions in a proposition.

Proposition 3.3 *The function* $\mathcal{C}(\tau) = \sum 1/\cos(\pi n \tau)$, *defined in the upper half-plane, satisfies*

(i) $\mathcal{C}(\tau + 2) = \mathcal{C}(\tau)$.

(ii) $\mathcal{C}(\tau) = (i/\tau)\mathcal{C}(-1/\tau)$.

---

[4] Why we refer to the point $\tau = 1$ as a cusp, and the reason for its importance, will become clear later on.

(iii) $\mathcal{C}(\tau) \to 1$ *as* $\mathrm{Im}(\tau) \to \infty$.

(iv) $\mathcal{C}(1 - 1/\tau) \sim 4(\tau/i)e^{\pi i\tau/2}$ *as* $\mathrm{Im}(\tau) \to \infty$.

Moreover, $\theta(\tau)^2$ *satisfies the same properties.*

With this proposition, we prove the identity of $\theta(\tau)^2 = \mathcal{C}(\tau)$ with the aid of the following theorem, in which we shall ultimately set $f = \mathcal{C}/\theta^2$.

Theorem 3.4 *Suppose* f *is a holomorphic function in the upper half-plane that satisfies:*

(i) $f(\tau + 2) = f(\tau)$,

(ii) $f(-1/\tau) = f(\tau)$,

(iii) $f(\tau)$ *is bounded,*

then f *is constant.*

For the proof of this theorem, we introduce the following subset of the closed upper half-plane, which is defined by

$$\mathcal{F} = \{\tau \in \overline{\mathbb{H}} : |\mathrm{Re}(\tau)| \leq 1 \text{ and } |\tau| \geq 1\},$$

and illustrated in Figure 1.

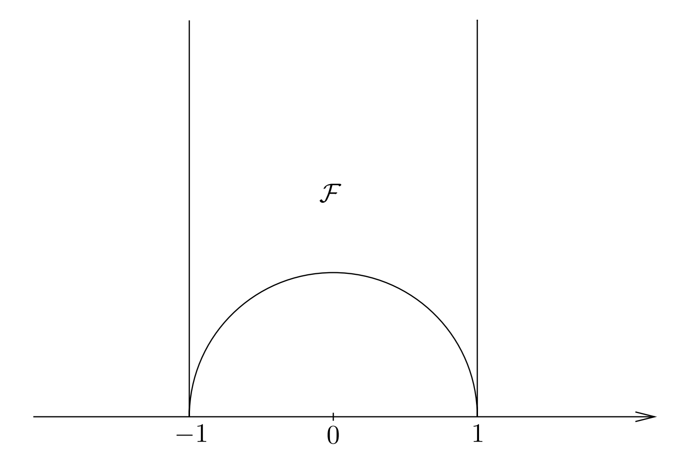


Figure 1. The domain $\mathcal{F}$

The points corresponding to $\tau = \pm 1$ are called **cusps**. They are equivalent under the mapping $\tau \mapsto \tau + 2$.

Lemma 3.5 *Every point in the upper half-plane can be mapped into $\mathcal{F}$ using repeatedly one or another of the following fractional linear transformations or their inverses:*

$$T_2 : \tau \mapsto \tau + 2, \qquad S : \tau \mapsto -1/\tau.$$

For this reason, $\mathcal{F}$ is called the **fundamental domain**[5] for the group of transformations generated by T_2 and S.

In fact, we let G denote the group generated by T_2 and S. Since T_2 and S are fractional linear transformations, we may represent an element $g \in G$ by a matrix

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with the understanding that

$$g(\tau) = \frac{a\tau + b}{c\tau + d}.$$

Since the matrices representing T_2 and S have integer coefficients and determinant 1, the same is true for all matrices of elements in G. In particular, if $\tau \in \mathbb{H}$, then

$$\text{(9)} \qquad \operatorname{Im}(g(\tau)) = \frac{\operatorname{Im}(\tau)}{|c\tau + d|^2}.$$

Proof of Lemma 3.5. Let $\tau \in \mathbb{H}$. If $g \in G$ with $g(\tau) = (a\tau + b)/(c\tau + d)$, then c and d are integers, and by (9) we may choose a $g_0 \in G$ such that $\operatorname{Im}(g_0(\tau))$ is maximal. Since the translations T_2 and their inverses do not change imaginary parts, we may apply finitely many of them to see that there exists $g_1 \in G$ with $|\operatorname{Re}(g_1(\tau))| \leq 1$ and $\operatorname{Im}(g_1(\tau))$ is maximal. It now suffices to prove that $|g_1(\tau)| \geq 1$ to conclude that $g_1(\tau) \in \mathcal{F}$. If this were not true, that is, $|g_1(\tau)| < 1$, then $\operatorname{Im}(Sg_1(\tau))$ would be greater than $\operatorname{Im}(g_1(\tau))$ since

$$\operatorname{Im}(Sg_1(\tau)) = \operatorname{Im}(-1/g_1(\tau)) = -\frac{\operatorname{Im}(\overline{g_1(\tau)})}{|g_1(\tau)|^2} > \operatorname{Im}(g_1(\tau)),$$

and this contradicts the maximality of $\operatorname{Im}(g_1(\tau))$.

We can now prove the theorem. Suppose f is not constant, and let $g(z) = f(\tau)$ where $z = e^{\pi i \tau}$. The function g is well defined for z in the

[5]Strictly speaking, the notion of a fundamental domain requires that every point have a unique representative in the domain. In the present case, ambiguity arises only for points that are on the boundary of $\mathcal{F}$.

punctured unit disc, since f is periodic of period 2, and moreover, g is bounded near the origin by assumption (iii) of the theorem. Hence 0 is a removable singularity for g, and $\lim_{z\to 0} g(z) = \lim_{\mathrm{Im}(\tau)\to\infty} f(\tau)$ exists. So by the maximum modulus principle,

$$\lim_{\mathrm{Im}(\tau)\to\infty} |f(\tau)| < \sup_{\tau\in\mathcal{F}} |f(\tau)|.$$

Now we must investigate the behavior of f at the points $\tau = \pm 1$. Since $f(\tau+2) = f(\tau)$, it suffices to consider the point $\tau = 1$. We claim that

$$\lim_{\mathrm{Im}(\tau)\to\infty} f(1 - 1/\tau)$$

exists and moreover

$$\lim_{\mathrm{Im}(\tau)\to\infty} |f(1 - 1/\tau)| < \sup_{\tau\in\mathcal{F}} |f(\tau)|.$$

The argument is essentially the same as the one above, except that we first need to interchange $\tau = 1$ with the point at infinity. In other words, we wish to investigate the behavior of $F(\tau) = f(1 - 1/\tau)$ for τ near ∞. The important step is to prove that F is periodic. To this end, we consider the fractional linear transformation associated to the matrix

$$U_n = \begin{pmatrix} 1-n & n \\ -n & 1+n \end{pmatrix},$$

that is,

$$\tau \mapsto \frac{(1-n)\tau + n}{-n\tau + (1+n)},$$

which maps 1 to 1. Now let $\mu(\tau) = 1/(1-\tau)$ which maps 1 to ∞, and whose inverse $\mu^{-1}(\tau) = 1 - 1/\tau$ takes ∞ to 1. Then

$$U_n = \mu^{-1} T_n \mu,$$

where T_n is the translation $T_n(\tau) = \tau + n$. As a consequence,

$$U_n U_m = U_{n+m},$$

and

$$U_{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} = T_2 S.$$

Thus any U_n can be obtained by finitely many applications of T_2, S, or their inverses. Since f is invariant under T_2 and S, it is also invariant under U_m. So we find that

$$f(\mu^{-1}T_n\mu(\tau)) = f(\tau).$$

Therefore, if we let $F(\tau) = f(\mu^{-1}(\tau)) = f(1 - 1/\tau)$, we find that F is periodic of period 1, that is,

$$F(T_n\tau) = F(\tau) \quad \text{for every integer } n.$$

Now, by the previous argument, if we set $h(z) = F(\tau)$ with $z = e^{2\pi i\tau}$, we see that h has a removable singularity at $z = 0$, and the desired inequality follows by the maximum principle.

We conclude from this analysis that f attains its maximum in the interior of the upper half-plane, and this contradicts the maximum principle.

The proof of the two-squares theorem is now only one step away.

We consider the function $f(\tau) = \mathcal{C}(\tau)/\theta(\tau)^2$. Since we know by the product formula that $\theta(\tau)$ does not vanish in the upper half-plane (Corollary 1.4), we find that f is holomorphic in $\mathbb{H}$. Moreover, by Proposition 3.3, f is invariant under the transformations T_2 and S, that is, $f(\tau + 2) = f(\tau)$ and $f(-1/\tau) = f(\tau)$. Finally, in the fundamental domain $\mathcal{F}$, the function $f(\tau)$ is bounded, and in fact tends to 1 as $\text{Im}(\tau)$ tends to infinity, or as τ tends to the cusps ± 1. This is because of properties (iii) and (iv) in Proposition 3.3, which are verified by both $\mathcal{C}$ and θ^2. Thus f is bounded in $\mathbb{H}$. The result is that f is a constant, which must be 1, proving that $\theta(\tau)^2 = \mathcal{C}(\tau)$, and with it the two-squares theorem.

3.2 The four-squares theorem

Statement of the theorem

In the rest of this chapter, we shall consider the case of four squares. More precisely, we will prove that every positive integer is the sum of four squares, and moreover we will determine a formula for $r_4(n)$ that describes the number of ways this can be done.

We need to introduce another divisor function, which we denote by $\sigma_1^*(n)$, and which equals the sum of divisors of n that are not divisible by 4. The main theorem we shall prove is the following.

Theorem 3.6 *Every positive integer is the sum of four squares, and moreover*

$$r_4(n) = 8\sigma_1^*(n) \quad \textit{for all } n \geq 1.$$

As before, we relate the sequence $\{r_4(n)\}$ via its generating function to an appropriate power of the function θ, which in this case is its fourth power. The result is that

$$\theta(\tau)^4 = \sum_{n=0}^{\infty} r_4(n) q^n$$

whenever $q = e^{\pi i \tau}$ with $\tau \in \mathbb{H}$.

The next step is to find the modular function whose equality with $\theta(\tau)^4$ expresses the identity $r_4(n) = 8\sigma_1^*(n)$. Unfortunately, here there is nothing as simple as the function $\mathcal{C}(\tau)$ that arose in the two-squares theorem; instead we shall need to construct a rather subtle variant of the Eisenstein series considered in the previous chapter. In fact, we define

$$E_2^*(\tau) = \sum_m \sum_n \frac{1}{\left(\frac{m\tau}{2} + n\right)^2} - \sum_m \sum_n \frac{1}{\left(m\tau + \frac{n}{2}\right)^2}$$

for $\tau \in \mathbb{H}$. The indicated order of summation is critical, since the above series do not converge absolutely. The following reduces the four-squares theorem to the modular properties of E_2^*.

Proposition 3.7 *The assertion $r_4(n) = 8\sigma_1^*(n)$ is equivalent to the identity*

$$\theta(\tau)^4 = \frac{-1}{\pi^2} E_2^*(\tau), \qquad \textit{where } \tau \in \mathbb{H}.$$

Proof. It suffices to prove that if $q = e^{\pi i \tau}$, then

$$\frac{-1}{\pi^2} E_2^*(\tau) = 1 + \sum_{k=1}^{\infty} 8\sigma_1^*(k) q^k.$$

First, recall the forbidden Eisenstein series that we considered in the last section of the previous chapter, and which is defined by

$$F(\tau) = \sum_m \left[\sum_n \frac{1}{(m\tau + n)^2} \right],$$

where the term $n = m = 0$ is omitted. Since the sum above is not absolutely convergent, the order of summation, first in n and then in m, is crucial. With this in mind, the definitions of E_2^* and F give immediately

$$E_2^*(\tau) = F\left(\frac{\tau}{2}\right) - 4F(2\tau). \tag{10}$$

In Corollary 2.6 (and Exercise 7) of the last chapter, we proved that

$$F(\tau) = \frac{\pi^2}{3} - 8\pi^2 \sum_{k=1}^{\infty} \sigma_1(k) e^{2\pi i k\tau},$$

where $\sigma_1(k)$ is the sum of the divisors of k.

Now observe that

$$\sigma_1^*(n) = \begin{cases} \sigma_1(n) & \text{if } n \text{ is not divisible by 4,} \\ \sigma_1(n) - 4\sigma_1(n/4) & \text{if } n \text{ is divisible by 4.} \end{cases}$$

Indeed, if n is not divisible by 4, then no divisors of n are divisible by 4. If $n = 4\tilde{n}$, and d is a divisor of n that is divisible by 4, say $d = 4\tilde{d}$, then $\tilde{d}$ divides $\tilde{n}$. This gives the second formula. Therefore, from this observation and (10) we find that

$$E_2^*(\tau) = -\pi^2 - 8\pi^2 \sum_{k=1}^{\infty} \sigma_1^*(k) e^{\pi i k\tau},$$

and the proof of the proposition is complete.

We have therefore reduced Theorem 3.6 to the identity $\theta^4 = -\pi^{-2} E_2^*$, and the key to establish this relation is that E_2^* satisfies the same modular properties as $\theta(\tau)^4$.

Proposition 3.8 *The function $E_2^*(\tau)$ defined in the upper half-plane has the following properties:*

(i) $E_2^*(\tau + 2) = E_2^*(\tau)$.

(ii) $E_2^*(\tau) = -\tau^{-2} E_2^*(-1/\tau)$.

(iii) $E_2^*(\tau) \to -\pi^2$ *as* $\text{Im}(\tau) \to \infty$.

(iv) $|E_2^*(1 - 1/\tau)| = O(|\tau^2 e^{\pi i \tau}|)$ *as* $\text{Im}(\tau) \to \infty$.

Moreover $-\pi^2\theta^4$ has the same properties.

The periodicity (i) of E_2^* is immediate from the definition. The proofs of the other properties of E_2^* are a little more involved.

Consider the forbidden Eisenstein series F and its reverse $\tilde{F}$, which is obtained from reversing the order of summation:

$$F(\tau) = \sum_m \sum_n \frac{1}{(m\tau + n)^2} \quad \text{and} \quad \tilde{F}(\tau) = \sum_n \sum_m \frac{1}{(m\tau + n)^2}.$$

In both cases, the term $n = m = 0$ is omitted.

Lemma 3.9 *The functions F and $\tilde{F}$ satisfy:*

(a) $F(-1/\tau) = \tau^2\tilde{F}(\tau)$,

(b) $F(\tau) - \tilde{F}(\tau) = 2\pi i/\tau$,

(c) $F(-1/\tau) = \tau^2 F(\tau) - 2\pi i\tau$.

Proof. Property (a) follows directly from the definitions of F and $\tilde{F}$, and the identity

$$(n + m(-1/\tau))^2 = \tau^{-2}(-m + n\tau)^2.$$

To prove property (b), we invoke the functional equation for the Dedekind eta function which was established earlier:

$$\eta(-1/\tau) = \sqrt{\tau/i}\ \eta(\tau),$$

where $\eta(\tau) = q^{1/12}\prod_{n=1}^{\infty}(1 - q^{2n})$, and $q = e^{\pi i\tau}$.

First, we take the logarithmic derivative of η with respect to the variable τ to find (by Proposition 3.2 in Chapter 5)

$$(\eta'/\eta)(\tau) = \frac{\pi i}{12} - 2\pi i\sum_{n=1}^{\infty}\frac{nq^{2n}}{1 - q^{2n}}.$$

However, if $\sigma_1(k)$ denotes the sum of the divisors of k, then one sees that

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{nq^{2n}}{1 - q^{2n}} &= \sum_{n=1}^{\infty}\sum_{\ell=0}^{\infty} nq^{2n}q^{2\ell n}\\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} nq^{2nm}\\ &= \sum_{k=1}^{\infty}\sigma_1(k)q^{2k}.\end{aligned}$$

If we recall that $F(\tau) = \pi^2/3 - 8\pi^2\sum_{k=1}^{\infty}\sigma_1(k)q^{2k}$, we find

$$(\eta'/\eta)(\tau) = \frac{i}{4\pi}F(\tau).$$

By the chain rule, the logarithmic derivative of $\eta(-1/\tau)$ is $\tau^{-2}(\eta'/\eta)(-1/\tau)$, and using property (a), we see that the logarithmic derivative of $\eta(-1/\tau)$

equals $(i/4\pi)\tilde{F}(\tau)$. Therefore, taking the logarithmic derivative of the functional equation for η we find

$$\frac{i}{4\pi}\tilde{F}(\tau) = \frac{1}{2\tau} + \frac{i}{4\pi}F(\tau),$$

and this gives $\tilde{F}(\tau) = -2\pi i/\tau + F(\tau)$, as desired.

Finally, (c) is a consequence of (a) and (b).

To prove the transformation formula (ii) for E_2^* under $\tau \mapsto -1/\tau$, we begin with

$$E_2^*(\tau) = F(\tau/2) - 4F(2\tau).$$

Then

$$\begin{aligned} E_2^*(-1/\tau) &= F(-1/(2\tau)) - 4F(-2/\tau) \\ &= [4\tau^2 F(2\tau) - 4\pi i\tau] - 4[(\tau/2)^2 F(\tau/2) - \pi i\tau] \\ &= 4\tau^2 F(2\tau) - 4(\tau^2/4)F(\tau/2) \\ &= -\tau^2(F(\tau/2) - 4F(2\tau)) \\ &= -\tau^2 E_2^*(\tau), \end{aligned}$$

as desired. To prove the third property recall that

$$F(\tau) = \frac{\pi^2}{3} - 8\pi^2 \sum_{k=1}^{\infty} \sigma_1(k) e^{2\pi i k\tau},$$

where the sum goes to 0 as $\mathrm{Im}(\tau) \to \infty$. Then, if we use the fact that

$$E_2^*(\tau) = F(\tau/2) - 4F(2\tau),$$

we conclude that $E_2^*(\tau) \to -\pi^2$ as $\mathrm{Im}(\tau) \to \infty$.

To prove the final property, we begin by showing that

$$E_2^*(1 - 1/\tau) = \tau^2 \left[F\left(\frac{\tau - 1}{2}\right) - F(\tau/2) \right]. \tag{11}$$

From the transformation formulas for F we have

$$\begin{aligned} F(1/2 - 1/2\tau) &= F\left(\frac{\tau - 1}{2\tau}\right) \\ &= \left(\frac{2\tau}{\tau - 1}\right)^2 F\left(\frac{2\tau}{1 - \tau}\right) - 2\pi i\frac{2\tau}{1 - \tau}, \end{aligned}$$

and

$$\begin{aligned} F\left(\frac{2\tau}{1-\tau}\right) &= F(-2+2/(1-\tau)) \\ &= F(2/(1-\tau)) \\ &= \left(\frac{1-\tau}{2}\right)^2 F\left(\frac{\tau-1}{2}\right) - 2\pi i\left(\frac{\tau-1}{2}\right). \end{aligned}$$

Hence,

$$F(1/2-1/2\tau) = \tau^2 F\left(\frac{\tau-1}{2}\right) - \frac{2\pi i 2\tau}{1-\tau} - 2\pi i\frac{(2\tau)^2}{(\tau-1)^2}\left(\frac{\tau-1}{2}\right).$$

But $F(2-2/\tau) = F(-2/\tau) = (\tau^2/4)F(\tau/2) - 2\pi i\tau/2$, and hence

$$\begin{aligned} E_2^*(1-1/\tau) &= F(1/2-1/2\tau) - 4F(2-2/\tau) \\ &= \tau^2\left[F\left(\frac{\tau-1}{2}\right) - F(\tau/2)\right] - 2\pi i\left(\frac{2\tau}{1-\tau} + \frac{2\tau^2}{\tau-1}\right) + 4\pi i\tau \\ &= \tau^2\left[F\left(\frac{\tau-1}{2}\right) - F(\tau/2)\right]. \end{aligned}$$

This proves (11). Then, the last property follows from it and the fact that

$$F(\tau) = \frac{\pi^2}{3} - 8\pi^2\sum_{k=1}^{\infty}\sigma_1(k)e^{2\pi ik\tau}.$$

Thus Proposition 3.8 is proved.

We can now conclude the proof of the four-squares theorem by considering the quotient $f(\tau) = E_2^*(\tau)/\theta(\tau)^4$, and applying Theorem 3.4, as in the two-squares theorem. Recall $\theta(\tau)^4 \to 1$ and $\theta(1-1/\tau)^4 \sim 16\tau^2 e^{\pi i\tau}$, as $\mathrm{Im}(\tau) \to \infty$. The result is that $f(\tau)$ is a constant, which equals $-\pi^2$ by Proposition 3.8. This completes the proof of the four-squares theorem.

4 Exercises

1. Prove that

$$\frac{(\Theta'(z|\tau))^2 - \Theta(z|\tau)\Theta''(z|\tau)}{\Theta(z|\tau)^2} = \wp_\tau(z-1/2-\tau/2) + c_\tau,$$

where c_τ can be expressed in terms of the first two derivatives of $\Theta(z|\tau)$, with respect to z, at $z = 1/2+\tau/2$. Compare this formula with the result in Exercise 5 in the previous chapter.

2. Consider the **Fibonacci numbers** $\{F_n\}_{n=0}^{\infty}$, defined by the two initial values $F_0 = 0$, $F_1 = 1$ and the recursion relation

$$F_n = F_{n-1} + F_{n-2} \qquad \text{for } n \geq 2.$$

(a) Consider the generating function $F(x) = \sum_{n=0}^{\infty} F_n x^n$ associated to $\{F_n\}$, and prove that

$$F(x) = x^2 F(x) + xF(x) + x$$

for all x in a neighborhood of 0.

(b) Show that the polynomial $q(x) = 1 - x - x^2$ can be factored as

$$q(x) = (1 - \alpha x)(1 - \beta x),$$

where α and β are the roots of the polynomial $p(x) = x^2 - x - 1$.

(c) Expand the expression for F in partial fractions and obtain

$$F(x) = \frac{x}{1 - x - x^2} = \frac{x}{(1-\alpha x)(1-\beta x)} = \frac{A}{1-\alpha x} + \frac{B}{1-\beta x},$$

where $A = 1/(\alpha - \beta)$ and $B = 1/(\beta - \alpha)$.

(d) Conclude that $F_n = A\alpha^n + B\beta^n$ for $n \geq 0$. The two roots of p are actually

$$\alpha = \frac{1+\sqrt{5}}{2} \qquad \text{and} \qquad \beta = \frac{1-\sqrt{5}}{2},$$

so that $A = 1/\sqrt{5}$ and $B = -1/\sqrt{5}$.

The number $1/\alpha = (\sqrt{5} - 1)/2$, which is known as the **golden mean**, satisfies the following property: given a line segment $[AC]$ of unit length (Figure 2), there exists a unique point B on this segment so that the following proportion holds

$$\frac{AC}{AB} = \frac{AB}{BC}.$$

If $\ell = AB$, this reduces to the equation $\ell^2 + \ell - 1 = 0$, whose only positive solution is the golden mean. This ratio arises also in the construction of the regular pentagon. It has played a role in architecture and art, going back to the time of ancient Greece.

3. More generally, consider the difference equation given by the initial values u_0 and u_1, and the recurrence relation $u_n = au_{n-1} + bu_{n-2}$ for $n \geq 2$. Define the generating function associated to $\{u_n\}_{n=0}^{\infty}$ by $U(x) = \sum_{n=0}^{\infty} u_n x^n$. The recurrence relation implies that $U(x)(1 - ax - bx^2) = u_0 + (u_1 - au_0)x$ in a neighborhood of

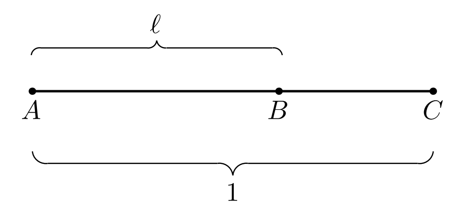


Figure 2. Appearance of the golden mean

the origin. If α and β denote the roots of the polynomial $p(x) = x^2 - ax - b$, then we may write

$$U(x) = \frac{u_0 + (u_1 - au_0)x}{(1-\alpha x)(1-\beta x)} = \frac{A}{1-\alpha x} + \frac{B}{(1-\beta x)} = A\sum_{n=0}^{\infty} \alpha^n x^n + B\sum_{n=0}^{\infty} \beta^n x^n,$$

where it is an easy matter to solve for A and B. Finally, this gives $u_n = A\alpha^n + B\beta^n$. Note that this approach yields a solution to our problem if the roots of p are distinct, namely $\alpha \neq \beta$. A variant of the formula holds if $\alpha = \beta$.

4. Using the generating formula for $p(n)$, prove the recurrence formula

$$\begin{aligned} p(n) &= p(n-1) + p(n-2) - p(n-5) - p(n-7) - \cdots \\ &= \sum_{k\neq 0} (-1)^{k+1} p\left(n - \frac{k(3k+1)}{2}\right), \end{aligned}$$

where the right-hand side is the finite sum taken over those $k \in \mathbb{Z}$, $k \neq 0$, with $k(3k+1)/2 \leq n$. Use this formula to calculate $p(5)$, $p(6)$, $p(7)$, $p(8)$, $p(9)$, and $p(10)$; check that $p(10) = 42$.

The next two exercises give elementary results related to the asymptotics of the partition function. More refined statements can be found in Appendix A.

5. Let

$$F(x) = \sum_{n=0}^{\infty} p(n)x^n = \prod_{n=1}^{\infty} \frac{1}{1-x^n}$$

be the generating function for the partitions. Show that

$$\log F(x) \sim \frac{\pi^2}{6(1-x)} \qquad \text{as } x \to 1, \text{ with } 0 < x < 1.$$

[Hint: Use $\log F(x) = \sum \log(1/(1-x^n))$ and $\log(1/(1-x^n)) = \sum (1/m)x^{nm}$, so

$$\log F(x) = \sum \frac{1}{m} \frac{x^m}{1-x^m}.$$

Use also $mx^{m-1}(1-x) < 1 - x^m < m(1-x)$.]

6. Show as a consequence of Exercise 5 that

$$e^{c_1 n^{1/2}} \leq p(n) \leq e^{c_2 n^{1/2}}$$

for two positive constants c_1 and c_2.
[Hint: $F(e^{-y}) = \sum p(n)e^{-ny} \leq Ce^{c/y}$ as $y \to 0$. So $p(n)e^{-ny} \leq ce^{c/y}$. Take $y = 1/n^{1/2}$ to get $p(n) \leq c' e^{c' n^{1/2}}$. In the opposite direction

$$\sum_{n=0}^{m} p(n)e^{-ny} \geq C(e^{c/y} - \sum_{n=m+1}^{\infty} e^{cn^{1/2}} e^{-ny}),$$

and it suffices to take $y = Am^{-1/2}$ where A is a large constant, and use the fact that the sequence $p(n)$ is increasing.]

7. Use the product formula for Θ to prove:

(a) The "triangular number" identity

$$\prod_{n=0}^{\infty} (1 + x^n)(1 - x^{2n+2}) = \sum_{n=-\infty}^{\infty} x^{n(n+1)/2},$$

which holds for $|x| < 1$.

(b) The "septagonal number" identity

$$\prod_{n=0}^{\infty} (1 - x^{5n+1})(1 - x^{5n+4})(1 - x^{5n+5}) = \sum_{n=-\infty}^{\infty} (-1)^n x^{n(5n+3)/2},$$

which holds for $|x| < 1$.

8. Consider Pythagorean triples (a, b, c) with $a^2 + b^2 = c^2$, and with $a, b, c \in \mathbb{Z}$. Suppose moreover that a and b have no common factors.

(a) Show that either a or b must be odd, and the other even.

(b) Show in this case (assuming a is odd and b even) that there are integers m, n so that $a = m^2 - n^2$, $b = 2mn$, and $c = m^2 + n^2$. [Hint: Note that $b^2 = (c - a)(c + a)$, and prove that $(c - a)/2$ and $(c + a)/2$ are relatively prime integers.]

(c) Conversely, show that whenever c is a sum of two-squares, then there exist integers a and b such that $a^2 + b^2 = c^2$.

9. Use the formula for $r_2(n)$ to prove the following:

(a) If $n = p$, where p is a prime of the form $4k + 1$, then $r_2(n) = 8$. This implies that n can be written in a unique way as $n = n_1^2 + n_2^2$, except for the signs and reordering of n_1 and n_2.

(b) If $n = q^a$, where q is prime of the form $4k + 3$ and a is a positive integer, then $r_2(n) > 0$ if and only if a is even.

(c) In general, n can be represented as the sum of two squares if and only if all the primes of the form $4k+3$ that arise in the prime decomposition of n occur with even exponents.

10. Observe the following irregularities of the functions $r_2(n)$ and $r_4(n)$ as n becomes large:

(a) $r_2(n) = 0$ for infinitely many n, while $\limsup_{n\to\infty} r_2(n) = \infty$.

(b) $r_4(n) = 24$ for infinitely many n while $\limsup_{n\to\infty} r_4(n)/n = \infty$.

[Hint: For (a) consider $n = 5^k$; for (b) consider alternatively $n = 2^k$, and $n = q^k$ where q is odd and large.]

11. Recall from Problem 2 in Chapter 2, that

$$\sum_{n=1}^{\infty} d(n)z^n = \sum_{n=1}^{\infty} \frac{z^n}{1-z^n}, \qquad |z| < 1$$

where $d(n)$ denotes the number of divisors of n.

More generally, show that

$$\sum_{n=1}^{\infty} \sigma_\ell(n)z^n = \sum_{n=1}^{\infty} \frac{n^\ell z^n}{1-z^n}, \qquad |z| < 1$$

where $\sigma_\ell(n)$ is the sum of the ℓth powers of divisors of n.

12. Here we give another identity involving θ^4, which is equivalent to the four-squares theorem.

(a) Show that for $|q| < 1$

$$\sum_{n=1}^{\infty} \frac{nq^n}{1-q^n} = \sum_{n=1}^{\infty} \frac{q^n}{(1-q^n)^2}.$$

[Hint: The left-hand side is $\sum \sigma_1(n)q^n$. Use $x/(1-x)^2 = \sum_{n=1}^{\infty} nx^n$.]

(b) Show as a result that

$$\sum_{n=1}^{\infty} \frac{nq^n}{1-q^n} - \sum_{n=1}^{\infty} \frac{4nq^{4n}}{1-q^{4n}} = \sum_{n=1}^{\infty} \frac{q^n}{(1-q^n)^2} - 4\sum_{n=1}^{\infty} \frac{q^{4n}}{(1-q^{4n})^2} = \sum \sigma_1^*(n)q^n$$

where $\sigma_1^*(n)$ is the sum of the divisors of d that are not divisible by 4.

(c) Show that the four-squares theorem is equivalent to the identity

$$\theta(\tau)^4 = 1 + 8\sum_{n=1}^{\infty} \frac{q^n}{(1+(-1)^n q^n)^2}, \qquad q = e^{\pi i \tau}.$$

5 Problems

1.* Suppose n is of the form $n = 4^a(8k+7)$, where a and k are positive integers. Show that n cannot be written as the sum of three-squares. The converse, that every n that is *not* of that form *can* be written as the sum of three-squares, is a difficult theorem of Legendre and Gauss.

2. Let $\mathrm{SL}_2(\mathbb{Z})$ denote the set of 2×2 matrices with integer entries and determinant 1, that is,

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathbb{Z} \ \text{ and } \ ad - bc = 1 \right\}.$$

This group acts on the upper half-plane by the fractional linear transformation $g(\tau) = (a\tau + b)/(c\tau + d)$. Together with this action comes the so-called fundamental domain $\mathcal{F}_1$ in the complex plane defined by

$$\mathcal{F}_1 = \{\tau \in \mathbb{C} : |\tau| \geq 1, \ \ |\mathrm{Re}(\tau)| \leq 1/2 \text{ and } |\mathrm{Im}(\tau)| \geq 0\}.$$

It is illustrated in Figure 3.

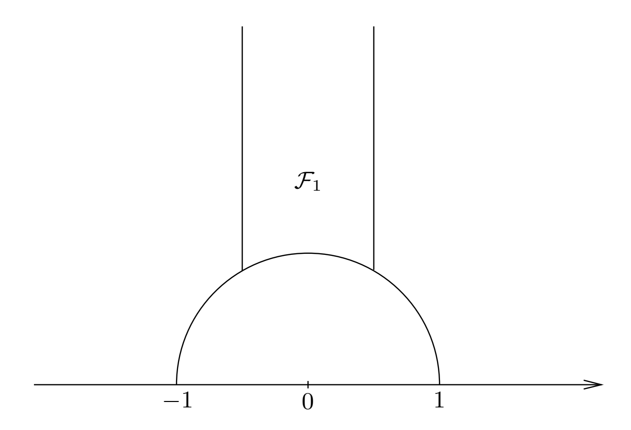


Figure 3. The fundamental domain $\mathcal{F}_1$

Consider the two elements in $\mathrm{SL}_2(\mathbb{Z})$ defined by $S(\tau) = -1/\tau$ and $T_1(\tau) = \tau + 1$. These correspond (for example) to the matrices

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

respectively. Let $\mathfrak{g}$ be the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by S and T_1.

(a) Show that for every $\tau \in \mathbb{H}$ there exists $g \in \mathfrak{g}$ such that $g(\tau) \in \mathcal{F}_1$.

(b) We say that two points τ and τ' are congruent if there exists $g \in \mathrm{SL}_2(\mathbb{Z})$ such that $g(\tau) = w$. Prove that if $\tau, w \in \mathcal{F}_1$ are congruent, then either $\mathrm{Re}(\tau) =$

$\pm 1/2$ and $\tau' = \tau \mp 1$ or $|\tau| = 1$ and $\tau' = -1/\tau$. [Hint: Say $\tau' = g(\tau)$. Why can one assume that $\text{Im}(\tau') \geq \text{Im}(\tau)$, and therefore $|c\tau + d| \leq 1$? Now consider separately the possibilities $c = -1$, $c = 0$, or $c = 1$.]

(c) Prove that S and T_1 generate the modular group in the sense that every fractional linear transformation corresponding to $g \in \text{SL}_2(\mathbb{Z})$ is a composition of finitely many S's and T_1's, and their inverses. Strictly speaking, the matrices associated to S and T_1 generate the projective special linear group $\text{PSL}_2(\mathbb{Z})$, which equals $\text{SL}_2(\mathbb{Z})$ modulo $\pm I$. [Hint: Observe that $2i$ is in the interior of $\mathcal{F}_1$. Now map $g(2i)$ back into $\mathcal{F}_1$ by using part (a). Use part (b) to conclude.]

3. In this problem, consider the group G of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer entries, determinant 1, and such that a and d have the same parity, b and c have the same parity, and c and d have opposite parity. This group also acts on the upper half-plane by fractional linear transformations. To the group G corresponds the fundamental domain $\mathcal{F}$ defined by $|\tau| \geq 1$, $|\text{Re}(\tau)| \leq 1$, and $\text{Im}(\tau) \geq 0$ (see Figure 1). Also, let

$$S(\tau) = -1/\tau \leftrightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad T_2(\tau) = \tau + 2 \leftrightarrow \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Prove that every fractional linear transformation corresponding to $g \in G$ is a composition of finitely many S, T_2 and their inverses, in analogy with the previous problem.

4. Let G denote the group of matrices given in the previous problem. Here we give an alternate proof of Theorem 3.4, that states that a function in $\mathbb{H}$ which is holomorphic, bounded, and invariant under G must be constant.

(a) Suppose that $f : \mathbb{H} \to \mathbb{C}$ is holomorphic, bounded, and that there exists a sequence of complex numbers $\tau_k = x_k + iy_k$ such that

$$f(\tau_k) = 0, \qquad \sum_{k=1}^{\infty} y_k = \infty, \qquad 0 < y_k \leq 1, \qquad \text{and} \qquad |x_k| \leq 1.$$

Then $f = 0$. [Hint: When $x_k = 0$ see Problem 5 in Chapter 8.]

(b) Given two relatively prime integers c and d with different parity, show that there exist integers a and b such that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$. [Hint: All the solutions of $xc + dy = 1$ take the form $x_0 + dt$ and $y_0 - ct$ where x_0, y_0 is a particular solution and $t \in \mathbb{Z}$.]

(c) Prove that $\sum 1/(c^2 + d^2) = \infty$ where the sum is taken over all c and d that are relatively prime and of opposite parity. [Hint: Suppose not, and prove that $\sum_{(a,b)=1} 1/(a^2 + b^2) < \infty$ where the sum is over all relatively prime integers a and b. To do so, note that if a and b are both odd and relatively prime, then the two numbers c and d defined by $c = (a + b)/2$

and $d = (a - b)/2$ are relatively prime and of opposite parity. Moreover, $c^2 + d^2 \le A(a^2 + b^2)$ for some universal constant A. Therefore

$$\sum_{n \neq 0} \frac{1}{n^2} \sum_{(a,b)=1} \frac{1}{a^2 + b^2} < \infty,$$

hence $\sum 1/(k^2 + \ell^2) < \infty$, where the sum is over all integers k and ℓ such that $k, \ell \neq 0$. Why is this a contradiction?]

(d) Prove that if $F : \mathbb{H} \to \mathbb{C}$ is holomorphic, bounded, and invariant under G, then F is constant. [Hint: Replace $F(\tau)$ by $F(\tau) - F(i)$ so that we can assume $F(i) = 0$ and prove $F = 0$. For each relatively prime c and d with opposite parity, choose $g \in G$ so that $g(i) = x_{c,d} + i/(c^2 + d^2)$ with $|x_{c,d}| \le 1$.]

5.* In Chapter 9 we proved that the Weierstrass $\wp$ function satisfies the cubic equation

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3,$$

where $g_2 = 60E_4$, $g_3 = 140E_6$, with E_k is the Eisenstein series of order k. The discriminant of the cubic $y^2 = 4x^3 - g_2 x - g_3$ is defined by $\triangle = g_2^3 - 27g_3^2$. Prove that

$$\triangle(\tau) = (2\pi)^{12}\eta^{24}(\tau) \qquad \text{for all } \tau \in \mathbb{H}.$$

[Hint: $\triangle$ and η^{24} satisfy the same transformation laws under $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$. Because of the fundamental domain described in Problem 2, it suffices then to investigate the behavior at the only cusp, which is at infinity.]

6.* Here we will deduce the formula for $r_8(n)$, which counts the number of representations of n as a sum of eight squares. The method is parallel to that of $r_4(n)$, but the details are less delicate.

Theorem: $r_8(n) = 16\sigma_3^*(n)$.

Here $\sigma_3^*(n) = \sigma_3(n) = \sum_{d|n} d^3$, when n is odd. Also, when n is even

$$\sigma_3^*(n) = \sum_{d|n} (-1)^d d^3 = \sigma_3^e(n) - \sigma_3^o(n),$$

where $\sigma_3^e(n) = \sum_{d|n,\ d \text{ even}} d^3$ and $\sigma_3^o(n) = \sum_{d|n,\ d \text{ odd}} d^3$.

Consider the appropriate Eisenstein series

$$E_4^*(\tau) = \sum \frac{1}{(n + m\tau)^4},$$

where the sum is over integers n and m with opposite parity. Recall the standard Eisenstein series

$$E_4(\tau) = \sum_{(n,m) \neq (0,0)} \frac{1}{(n + m\tau)^4}.$$

Notice that the series defining E_4^* is absolutely convergent, in distinction to $E_2^*(\tau)$, which arose when considering $r_4(n)$. This makes some of the considerations below quite a bit simpler.

(a) Prove that $r_8(n) = 16\sigma_3^*(n)$ is equivalent to the identity $\theta(\tau)^8 = 48\pi^{-4}E_4^*(\tau)$. [Hint: Use the fact that $E_4(\tau) = 2\zeta(4) + \frac{(2\pi)^4}{3}\sum_{k=1}^{\infty}\sigma_3(k)e^{2\pi ik\tau}$ and $\zeta(4) = \pi^4/90$.]

(b) Note that $E_4^*(\tau) = E_4(\tau) - 2^{-4}E_4((\tau-1)/2)$.

(c) $E_4^*(\tau+2) = E_4^*(\tau)$.

(d) $E_4^*(\tau) = \tau^{-4}E_4^*(-1/\tau)$.

(e) $(48/\pi^4)E_4^*(\tau) \to 1$ as $\tau \to \infty$.

(f) $|E_4^*(1-1/\tau)| \approx |\tau|^4|e^{2\pi i\tau}|$, as $\text{Im}(\tau) \to \infty$. [Hint: Verify that $E_4^*(1-1/\tau) = \tau^4(E_4(\tau) - E_4(2\tau))$.]

Since $\theta(\tau)^8$ satisfies properties similar to (c), (d), (e) and (f) above, it follows that the invariant function $48\pi^{-4}E_4^*(\tau)/\theta(\tau)^8$ is bounded and hence a constant, which must be 1. This gives the desired result.

Appendix A: Asymptotics

On the numerical computation of the definite integral $\int_w \cos\frac{\pi}{2}(w^3 - m.w)$, between the limits 0 and $\frac{1}{0}$.

The simplicity of the form of this differential coefficient induces me to suppose that the integral may possibly be expressible by some of the integrals whose values have been tabulated. After many attempts however, I have not succeeded in reducing it to any known integral: and I have therefore computed its value by actual summation to a considerable extent and by series for the remainder.

G. B. Airy, 1838

In a number of problems in analysis the solution is given by a function whose explicit calculation is not tractable. Often a useful substitute (and the only recourse) is to study the asymptotic behavior of this function near the point of interest. Here we shall investigate several related types of asymptotics, where the ideas of complex analysis are of crucial help. These typically center about the behavior for large values of the variable s of an integral of the form

$$I(s) = \int_a^b e^{-s\Phi(x)}\, dx. \tag{1}$$

We organize our presentation by formulating three guiding principles.

(i) **Deformation of contour**. The function Φ is in general complex-valued, therefore, for large s the integrand in (1) may oscillate rapidly, so that the resulting cancellations mask the true behavior of $I(s)$. When Φ is holomorphic (which is often the case) one can hope to change the contour of integration so that as far as possible, on the new contour Φ is essentially real-valued. If this is possible, one can then hope to read off the behavior of $I(s)$ in a rather direct manner. This idea will be illustrated first in the context of Bessel functions.

(ii) **Laplace's method**. In the case when Φ is real-valued on the contour and s is positive, the maximum contribution to $I(s)$ comes

from the integration near a minimum of Φ, and this leads to a satisfactory expansion in terms of the quadratic behavior of Φ near its minimum. We apply these ideas to present the asymptotics of the gamma function (Stirling's formula), and also those of the Airy function.

(iii) **Generating functions**. If $\{F_n\}$ is a number-theoretic or combinatorial sequence, we have already seen in several examples that one can exploit analytic properties of the generating function, $F(u) = \sum F_n u^n$, to obtain interesting conclusions regarding $\{F_n\}$. In fact the asymptotic behavior of F_n, as $n \to \infty$, can also be analyzed this way, via the formula

$$F_n = \int_\gamma F(e^{2\pi iz})e^{-2\pi inz}\, dz.$$

Here γ is an appropriate segment of unit length in the upper half-plane. This formula can then be studied as a variant of the integral (1). We shall show how these ideas apply in an important particular case to obtain an asymptotic formula for $p(n)$, the number of partitions of n.

1 Bessel functions

Bessel functions appear naturally in many problems that exhibit rotational symmetries. For instance, the Fourier transform of a spherical function in $\mathbb{R}^d$ is neatly expressed in terms of a Bessel function of order $(d/2) - 1$. See Chapter 6 in Book I.

The Bessel functions can be defined by a number of alternative formulas. We take the one that is valid for all order $\nu > -1/2$, given by

$$J_\nu(s) = \frac{(s/2)^\nu}{\Gamma(\nu + 1/2)\Gamma(1/2)} \int_{-1}^{1} e^{isx}(1 - x^2)^{\nu - 1/2}\, dx. \tag{2}$$

If we also write $J_{-1/2}(s)$ for $\lim_{\nu \to -1/2} J_\nu(s)$, we see that it equals $\sqrt{\frac{2}{\pi s}} \cos s$; observe in addition that $J_{1/2}(s) = \sqrt{\frac{2}{\pi s}} \sin s$. However, $J_\nu(s)$ has an expression in terms of elementary functions only when ν is half-integral, and understanding this function in general requires further analysis. Its behavior for large s is suggested by the two examples above.

Theorem 1.1 *$J_\nu(s) = \sqrt{\frac{2}{\pi s}} \cos\left(s - \frac{\pi\nu}{2} - \frac{\pi}{4}\right) + O\left(s^{-3/2}\right)$ as $s \to \infty$.*

In view of the formula for $J_\nu(s)$ it suffices to investigate

$$I(s) = \int_{-1}^{1} e^{isx}(1-x^2)^{\nu-1/2}\,dx, \tag{3}$$

and to this end we consider the analytic function $f(z) = e^{isz}(1-z^2)^{\nu-1/2}$ in the complex plane slit along the rays $(-\infty,-1)\cup(1,\infty)$; for $(1-z^2)^{\nu-1/2}$ we choose that branch that is positive when $z = x \in (-1,1)$. With $s > 0$ fixed, we apply Cauchy's theorem to see that

$$I(s) = -I_-(s) - I_+(s),$$

where the integrals $I(s)$, $I_-(s)$, and $I_+(s)$ are taken over the lines shown in Figure 1. This is established by using the fact that $\int_{\gamma_{\epsilon,R}} f(z)\,dz = 0$ where $\gamma_{\epsilon,R}$ is the second contour of Figure 1, and letting $\epsilon \to 0$ and $R \to \infty$.

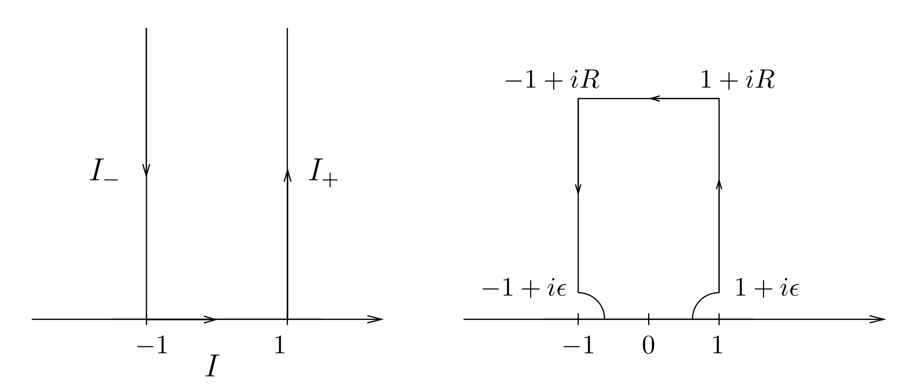


Figure 1. Contours of integration of $I(s)$, $I_-(s)$, $I_+(s)$, and the contour $\gamma_{\epsilon,R}$

On the contour for $I_+(s)$ we have $z = 1+iy$, so

$$I_+(s) = ie^{is}\int_0^\infty e^{-sy}(1-(1+iy)^2)^{\nu-1/2}\,dy. \tag{4}$$

There is a similar expression for $I_-(s)$.

What has the passage from $I(s)$ to $-(I_-(s)+I_+(s))$ gained us? Observe that for large positive s, the exponential e^{isx} in (3) oscillates rapidly, so the estimation of that integral is not obvious at first glance. However, in (4) the corresponding exponential is e^{-sy}, and it decreases

rapidly as $s \to \infty$, except when $y = 0$. Thus in this case one sees immediately that the main contribution to the integral comes from the integration near $y = 0$, and this allows one readily to approximate this integral. This idea is made precise in the following observation.

Proposition 1.2 *Suppose a and m are fixed, with $a > 0$ and $m > -1$. Then as $s \to \infty$*

$$\int_0^a e^{-sx} x^m \, dx = s^{-m-1}\Gamma(m+1) + O(e^{-cs}), \tag{5}$$

for some positive c.

Proof. The fact that $m > -1$ guarantees that $\int_0^a e^{-sx} x^m \, dx = \lim_{\epsilon \to 0} \int_\epsilon^a e^{-sx} x^m \, dx$ exists. Then, we write

$$\int_0^a e^{-sx} x^m \, dx = \int_0^\infty e^{-sx} x^m \, dx - \int_a^\infty e^{-sx} x^m \, dx.$$

The first integral on the right-hand side can be seen to equal $s^{-m-1}\Gamma(m+1)$, if we make the change of variables $x \mapsto x/s$. For the second integral we note that

$$\int_a^\infty e^{-sx} x^m \, dx = e^{-cs} \int_a^\infty e^{-s(x-c)} x^m \, dx = O(e^{-cs}), \tag{6}$$

as long as $c < a$, and so the proposition is proved.

We return to the integral (4) and observe that

$$(1 - (1+iy)^2)^{\nu-1/2} = (-2iy)^{\nu-1/2} + O(y^{\nu+1/2}), \quad \text{for } 0 \leq y \leq 1,$$

while

$$(1 - (1+iy)^2)^{\nu-1/2} = O(y^{\nu-1/2} + y^{2\nu-1}), \quad \text{for } 1 \leq y.$$

So, applying the proposition with $a = 1$ and $m = \nu \mp 1/2$, as well as (6), gives

$$I_+(s) = i(-2i)^{\nu-1/2} e^{is} s^{-\nu-1/2}\Gamma(\nu + 1/2) + O(s^{-\nu-3/2}).$$

Similarly,

$$I_-(s) = i(2i)^{\nu-1/2} e^{is} s^{-\nu-1/2}\Gamma(\nu + 1/2) + O(s^{-\nu-3/2}).$$

If we recall that

$$J_\nu(s) = \frac{(s/2)^\nu}{\Gamma(\nu+1/2)\Gamma(1/2)}[-I_-(s) - I_+(s)],$$

and the fact that $\Gamma(1/2) = \sqrt{\pi}$, we see that we have obtained the proof of the theorem.

For later purposes it is interesting to point out that under certain restricted circumstances, the gist of the conclusion in Proposition 1.2 extends to the complex half-plane $\text{Re}(s) \geq 0$.

Proposition 1.3 *Suppose a and m are fixed, with $a > 0$ and $-1 < m < 0$. Then as $|s| \to \infty$ with $\text{Re}(s) \geq 0$,*

$$\int_0^a e^{-sx} x^m \, dx = s^{-m-1}\Gamma(m+1) + O(1/|s|).$$

(Here s^{-m-1} is the branch of that function that is positive for $s > 0$).

Proof. We begin by showing that when $\text{Re}(s) \geq 0$, $s \neq 0$,

$$\int_0^\infty e^{-sx} x^m \, dx = \lim_{N\to\infty} \int_0^N e^{-sx} x^m \, dx$$

exists and equals $s^{-m-1}\Gamma(m+1)$. If N is large, we first write

$$\int_0^N e^{-sx} x^m \, dx = \int_0^a e^{-sx} x^m \, dx + \int_a^N e^{-sx} x^m \, dx.$$

Since $m > -1$, the first integral on the right-hand side defines an analytic function everywhere. For the second integral, we note that $-\frac{1}{s}\frac{d}{dx}(e^{-sx}) = e^{-sx}$, so an integration by parts gives

$$\int_a^N e^{-sx} x^m \, dx = \frac{m}{s} \int_a^N e^{-sx} x^{m-1} \, dx - \left[\frac{e^{-sx}}{s} x^m\right]_a^N. \tag{7}$$

This identity, together with the convergence of the integral $\int_a^\infty x^{m-1} dx$, shows that $\int_a^\infty e^{-sx} x^m \, dx$ defines an analytic function on $\text{Re}(s) > 0$ that is continuous on $\text{Re}(s) \geq 0$, $s \neq 0$. Thus $\int_0^\infty e^{-sx} x^m \, dx$ is analytic on the half-plane $\text{Re}(s) > 0$ and continuous on $\text{Re}(s) \geq 0$, $s \neq 0$. Since it equals $s^{-m-1}\Gamma(m+1)$ when s is positive, we deduce that $\int_0^\infty e^{-sx} x^m \, dx = s^{-m-1}\Gamma(m+1)$ when $\text{Re}(s) \geq 0$, $s \neq 0$.

However, we now have

$$\int_0^a e^{-sx} x^m \, dx = \int_0^\infty e^{-sx} x^m \, dx - \int_a^\infty e^{-sx} x^m \, dx.$$

It is clear from (7), and from the fact that $m < 0$, that if we let $N \to \infty$, then $\int_a^\infty e^{-sx} x^{m-1} \, dx = O(1/|s|)$. The proposition if therefore proved.

Note. If one wants to obtain a better error term in Proposition 1.3, or for that matter extend the range of m, then one needs to mitigate the effect of the contribution of the end-point $x = a$. This can be done by introducing suitable smooth cut-offs. See Problem 1.

2 Laplace's method; Stirling's formula

We have already mentioned that when Φ is real-valued, the main contribution to $\int_a^b e^{-s\Phi(x)} \, dx$ as $s \to \infty$ comes from the point where Φ takes its minimum value. A situation where this minimum is attained at one of the end-points, a or b, was considered in Proposition 1.2. We now turn to the important case when the minimum is achieved in the interior of $[a, b]$.

Consider

$$\int_a^b e^{-s\Phi(x)} \psi(x) \, dx$$

where the **phase** Φ is real-valued, and both it and the **amplitude** ψ are assumed for simplicity to be indefinitely differentiable. Our hypothesis regarding the minimum of Φ is that there is an $x_0 \in (a, b)$ so that $\Phi'(x_0) = 0$, but $\Phi''(x_0) > 0$ throughout $[a, b]$ (Figure 2 illustrates the situation.)

Proposition 2.1 *Under the above assumptions, with $s > 0$ and $s \to \infty$,*

$$\int_a^b e^{-s\Phi(x)} \psi(x) \, dx = e^{-s\Phi(x_0)} \left[\frac{A}{s^{1/2}} + O\left(\frac{1}{s}\right) \right], \tag{8}$$

where

$$A = \sqrt{2\pi} \frac{\psi(x_0)}{(\Phi''(x_0))^{1/2}}.$$

Proof. By replacing $\Phi(x)$ by $\Phi(x) - \Phi(x_0)$ we may assume that $\Phi(x_0) = 0$. Since $\Phi'(x_0) = 0$, we note that

$$\frac{\Phi(x)}{(x - x_0)^2} = \frac{\Phi''(x_0)}{2} \varphi(x),$$

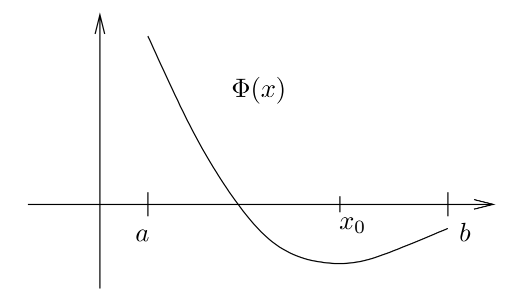


Figure 2. The function Φ, with its minimum at x_0

where φ is smooth, and $\varphi(x) = 1 + O(x - x_0)$ as $x \to x_0$. We can therefore make the smooth change of variables $x \mapsto y = (x - x_0)(\varphi(x))^{1/2}$ in a small neighborhood of $x = x_0$, and observe that $dy/dx|_{x_0} = 1$, and thus $dx/dy = 1 + O(y)$ as $y \to 0$. Moreover, we have $\psi(x) = \tilde{\psi}(y)$ with $\tilde{\psi}(y) = \psi(x_0) + O(y)$ as $y \to 0$. Hence if $[a', b']$ is a sufficiently small interval containing x_0 in its interior, by making the indicated change of variables we obtain

(9)
$$\int_{a'}^{b'} e^{-s\Phi(x)}\psi(x)\,dx = \psi(x_0)\int_{\alpha}^{\beta} e^{-s\frac{\Phi''(x_0)}{2}y^2}\,dy + O\left(\int_{\alpha}^{\beta} e^{-s\frac{\Phi''(x_0)}{2}y^2}|y|dy\right),$$

where $\alpha < 0 < \beta$. We now make the further change of variables $y^2 = X$, $dy = \frac{1}{2}X^{-1/2}\,dX$, and we see by (5) that the first integral on the right-hand side in (9) is

$$\int_0^{a_0} e^{-s\frac{\Phi''(x_0)}{2}X} X^{-1/2}\,dX + O(e^{-\delta s}) = s^{-1/2}\left(\frac{2\pi}{\Phi''(x_0)}\right)^{1/2} + O(e^{-\delta s}),$$

for some $\delta > 0$. By the same argument, the second integral is $O(1/s)$. What remains are the integrals of $e^{-s\Phi(x)}\psi(x)$ over $[a, a']$ and $[b', b]$; but these integrals decay exponentially as $s \to \infty$, since $\Phi(x) \geq c > 0$ in these two sub-intervals. Altogether, this establishes (8) and the proposition.

It is important to realize that the asymptotic relation (8) extends to all complex s with $\mathrm{Re}(s) \geq 0$. The proof, however, requires a somewhat different argument: here we must take into account the oscillations of $e^{-s\Phi(x)}$ when $|s|$ is large but $\mathrm{Re}(s)$ is small, and this is achieved by a simple integration by parts.

Proposition 2.2 *With the same assumptions on* Φ *and* ψ*, the relation* (8) *continues to hold if* $|s| \to \infty$ *with* $\mathrm{Re}(s) \geq 0$.

Proof. We proceed as before to the equation (9), and obtain the appropriate asymptotic for the first term, by virtue of Proposition 1.3, with $m = -1/2$. To deal with the rest we start with an observation. If Ψ and ψ are given on an interval $[\overline{a}, \overline{b}]$, are indefinitely differentiable, and $\Psi(x) \geq 0$, while $|\Psi'(x)| \geq c > 0$, then if $\mathrm{Re}(s) \geq 0$,

$$\int_{\overline{a}}^{\overline{b}} e^{-s\Psi(x)} \psi(x)\, dx = O\left(\frac{1}{|s|}\right) \qquad \text{as } |s| \to \infty. \tag{10}$$

Indeed, the integral equals

$$-\frac{1}{s} \int_{\overline{a}}^{\overline{b}} \frac{d}{dx}\left(e^{-s\Psi(x)}\right) \frac{\psi(x)}{\Psi'(x)}\, dx,$$

which by integration by parts gives

$$\frac{1}{s} \int_{\overline{a}}^{\overline{b}} e^{-s\Psi(x)} \frac{d}{dx}\left(\frac{\psi(x)}{\Psi'(x)}\right) dx - \frac{1}{s}\left[e^{-s\Psi(x)} \frac{\psi(x)}{\Psi'(x)}\right]_{\overline{a}}^{\overline{b}}.$$

The assertion (10) follows immediately since $|e^{-s\Psi(x)}| \leq 1$, when $\mathrm{Re}(s) \geq 0$. This allows us to deal with the integrals of $e^{-s\Phi(x)}\psi(x)$ in the complementary intervals $[a, a']$ and $[b', b]$, because in each, $|\Phi'(x)| \geq c > 0$, since $\Phi'(x_0) = 0$ and $\Phi''(x) \geq c_1 > 0$.

Finally, for the second term on the right-hand side of (9) we observe that it is actually of the form

$$\int_{\alpha}^{\beta} e^{-s\frac{\Phi''(x_0)}{2}y^2} y\eta(y)\, dy,$$

where $\eta(y)$ is differentiable. Then we can again estimate this term by integration by parts, once we write it as

$$-\frac{1}{s\Phi''(x_0)} \int_{\alpha}^{\beta} \frac{d}{dy}\left(e^{-s\frac{\Phi''(x_0)}{2}y^2}\right) \eta(y)\, dy,$$

obtaining the bound $O(1/|s|)$.

The special case of Proposition 2.2 when s is purely imaginary, $s = it$, $t \to \pm\infty$, is often treated separately; the argument in this situation is

usually referred to as the method of **stationary phase**. The points x_0 for which $\Phi'(x_0) = 0$ are called the **critical points**.

Our first application will be to the asymptotic behavior of the gamma function Γ, given by Stirling's formula. This formula will be valid in any sector of the complex plane that omits the negative real axis. For any $\delta > 0$ we set $S_\delta = \{s : |\arg s| \leq \pi - \delta\}$, and denote by $\log s$ the principal branch of the logarithm that is given in the plane slit along the negative real axis.

Theorem 2.3 *If* $|s| \to \infty$ *with* $s \in S_\delta$*, then*

$$\Gamma(s) = e^{s \log s} e^{-s} \frac{\sqrt{2\pi}}{s^{1/2}} \left(1 + O\left(\frac{1}{|s|^{1/2}}\right)\right). \tag{11}$$

Remark. With a little extra effort one can improve the error term to $O(1/|s|)$, and in fact obtain a complete asymptotic expansion in powers of $1/s$; see Problem 2. Also, we note that (11) implies $\Gamma(s) \sim \sqrt{2\pi} s^{s-1/2} e^{-s}$, which is how Stirling's formula is often stated.

To prove the theorem we first establish (11) in the right half-plane. We shall show that the formula holds whenever $\text{Re}(s) > 0$, and in addition that the error term is uniform on the closed half-plane, once we omit a neighborhood of the origin (say $|s| < 1$). To see this, start with $s > 0$, and write

$$\Gamma(s) = \int_0^\infty e^{-x} x^s \, \frac{dx}{x} = \int_0^\infty e^{-x + s \log x} \, \frac{dx}{x}.$$

Upon making the change of variables $x \mapsto sx$, the above equals

$$\int_0^\infty e^{-sx + s \log sx} \, \frac{dx}{x} = e^{s \log s} e^{-s} \int_0^\infty e^{-s\Phi(x)} \, \frac{dx}{x},$$

where $\Phi(x) = x - 1 - \log x$. By analytic continuation this identity continues to hold, and we have when $\text{Re}(s) > 0$,

$$\Gamma(s) = e^{s \log s} e^{-s} I(s)$$

with

$$I(s) = \int_0^\infty e^{-s\Phi(x)} \, \frac{dx}{x}.$$

It now suffices to see that

$$I(s) = \frac{\sqrt{2\pi}}{s^{1/2}} + O\left(\frac{1}{|s|}\right) \qquad \text{for} \qquad \text{Re}(s) > 0. \tag{12}$$

Observe first that $\Phi(1) = \Phi'(1) = 0$, $\Phi''(x) = 1/x^2 > 0$ whenever $0 < x < \infty$, and $\Phi''(1) = 1$. Thus Φ is convex, attains its minimum at $x = 1$, and is positive.

We apply the complex version of the Laplace method, Proposition 2.2, in this situation. Here the critical point is $x_0 = 1$ and $\psi(x) = 1/x$. We choose for convenience the interval $[a, b]$ to be $[1/2, 2]$. Then for $\int_a^b e^{-s\Phi(x)}\psi(x)\,dx$ we get the asymptotic (12). It remains to bound the error terms, those corresponding to integration over $[0, 1/2]$, and $[2, \infty)$. Here the device of integration by parts, which has served us so well, can be applied again. Indeed, since $\Phi'(x) = 1 - 1/x$, we have

$$\begin{aligned}\int_\epsilon^{1/2} e^{-s\Phi(x)}\,\frac{dx}{x} &= -\frac{1}{s}\int_\epsilon^{1/2} \frac{d}{dx}\left(e^{-s\Phi(x)}\right)\frac{dx}{\Phi'(x)x}\\ &= -\frac{1}{s}\left[\frac{e^{-s\Phi(x)}}{x-1}\right]_\epsilon^{1/2} - \frac{1}{s}\int_\epsilon^{1/2} e^{-s\Phi(x)}\frac{dx}{(x-1)^2}.\end{aligned}$$

Noting that $\Phi(\epsilon) \to +\infty$ as $\epsilon \to 0$, and $|e^{-s\Phi(x)}| \leq 1$, we find in the limit that

$$\int_0^{1/2} e^{-s\Phi(x)}\,\frac{dx}{x} = \frac{2}{s}e^{-s\Phi(1/2)} - \frac{1}{s}\int_0^{1/2} e^{-s\Phi(x)}\frac{dx}{(x-1)^2}.$$

Thus the left-hand side is $O(1/|s|)$ in the half-plane $\mathrm{Re}(s) \geq 0$.

The integral $\int_2^\infty e^{-s\Phi(x)}\,\frac{dx}{x}$ is treated analogously, once we note that $\int_2^\infty (x-1)^{-2}\,dx$ converges.

Since these estimates are uniform, (12) and thus (11) are proved for $\mathrm{Re}(s) \geq 0$, $|s| \to \infty$.

To pass from $\mathrm{Re}(s) \geq 0$ to $\mathrm{Re}(s) \leq 0$, $s \in S_\delta$, we record the following fact about the principal branch of $\log s$: whenever $\mathrm{Re}(s) \geq 0$, $s = \sigma + it$, $t \neq 0$, then

$$\log(-s) = \begin{cases} \log s - i\pi & \text{if } t > 0,\\ \log s + i\pi & \text{if } t < 0.\end{cases}$$

Hence if $G(s) = e^{s\log s}e^{-s}$, $\mathrm{Re}(s) \geq 0$, $t \neq 0$, then

$$G(-s)^{-1} = \begin{cases} e^{s\log s}e^{-s}e^{-si\pi} & \text{if } t > 0,\\ e^{s\log s}e^{-s}e^{si\pi} & \text{if } t < 0.\end{cases} \tag{13}$$

Next,

$$\Gamma(s)\Gamma(-s) = \frac{\pi}{-s\sin\pi s}, \tag{14}$$

which follows from the fact that $\Gamma(s)\Gamma(1-s) = \pi/\sin \pi s$, and $\Gamma(1-s) = -s\Gamma(-s)$ (see Theorem 1.4 and Lemma 1.2 in Chapter 6). The combination of (13) and (14), together with the fact that for large s, $\left(1 + O(1/|s|^{1/2})\right)^{-1} = 1 + O(1/|s|^{1/2})$, allows us then to extend (11) to the whole sector S_δ, thereby completing the proof of the theorem.

3 The Airy function

The Airy function appeared first in optics, and more precisely, in the analysis of the intensity of light near a caustic; it was an important early instance in the study of asymptotics of integrals, and it continues to arise in a number of other problems. The **Airy function** Ai is defined by

$$\text{(15)} \qquad \mathrm{Ai}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(x^3/3+sx)}\, dx, \qquad \text{with } s \in \mathbb{R}.$$

Let us first see that because of the rapid oscillations of the integrand as $|x| \to \infty$, the integral converges and represents a continuous function of s. In fact, note that

$$\frac{1}{i(x^2+s)} \frac{d}{dx} \left(e^{i(x^3/3+sx)} \right) = e^{i(x^3/3+sx)},$$

so if $a \geq 2|s|^{1/2}$, we can write the integral $\int_a^R e^{i(x^3/3+sx)}\, dx$ as

$$\text{(16)} \qquad \int_a^R \frac{1}{i(x^2+s)} \frac{d}{dx} \left(e^{i(x^3/3+sx)} \right) dx.$$

We may now integrate by parts and let $R \to \infty$, to see that the integral converges uniformly, and that as a result $\int_a^\infty e^{i(x^3/3+sx)}\, dx$ is also continuous for $|s| \leq a^2/4$. The same argument works for the integral from $-\infty$ to $-a$ and our assertion regarding $\mathrm{Ai}(s)$ is established.

A better insight into $\mathrm{Ai}(s)$ is given by deforming the contour of integration in (15). A choice of an optimal contour will appear below, but for now let us notice that as soon as we replace the x-axis of integration in (15) by the parallel line $L_\delta = \{x + i\delta,\ x \in \mathbb{R}\}$, $\delta > 0$, matters improve dramatically.

In fact, we may apply the Cauchy theorem to $f(z) = e^{i(z^3/3+sz)}$ over the rectangle γ_R shown in Figure 3.

One observes that $f(z) = O(e^{-\delta x^2})$ on L_δ, while $f(z) = O(e^{-yR^2})$ on the vertical sides of the rectangle. Thus since $\int_0^\delta e^{-yR^2}\, dy \to 0$ as

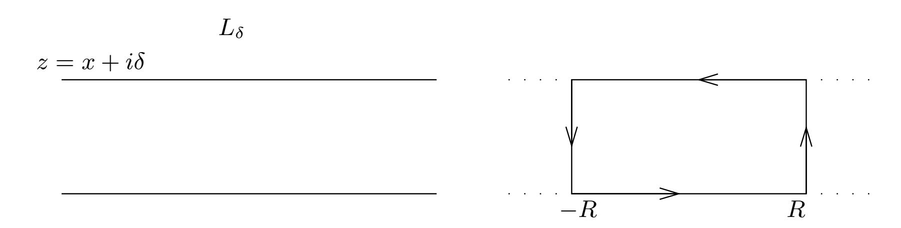


Figure 3. The line L_δ and the contour γ_R

$R \to \infty$, we see that

$$\mathrm{Ai}(s) = \frac{1}{2\pi} \int_{L_\delta} e^{i(z^3/3+sz)}\, dz.$$

Now the majorization $f(z) = O(e^{-\delta x^2})$ continues to hold for each complex s, and hence because of the (rapid) convergence of the integral, $\mathrm{Ai}(s)$ extends to an entire function of s.

We note next that $\mathrm{Ai}(s)$ satisfies the differential equation

$$\mathrm{Ai}''(s) = s\mathrm{Ai}(s). \tag{17}$$

This simple and natural equation helps to explain the ubiquity of the Airy function. To prove (17) observe that

$$\mathrm{Ai}''(s) - s\mathrm{Ai}(s) = \frac{1}{2\pi} \int_{L_\delta} (-z^2 - s)e^{i(z^3/3+sz)}\, dz.$$

But $-(z^2 + s)e^{i(z^3/3+sz)} = i\frac{d}{dz}(e^{i(z^3/3+sz)})$, so

$$\mathrm{Ai}''(s) - s\mathrm{Ai}(s) = \frac{i}{2\pi} \int_{L_\delta} \frac{d}{dz}(f(z))\, dz = 0,$$

since $f(z) = e^{i(z^3/3+sz)}$ vanishes as $|z| \to \infty$ along L_δ.

We now turn to our main problem, the asymptotics of $\mathrm{Ai}(s)$ for large (real) values of s. The differential equation (17) shows us that we may expect different behaviors of the Airy function when $|s|$ is large, depending on whether s is positive or negative. To see this, we compare the equation with a simple analogue

$$y''(s) = Ay(s), \tag{18}$$

where A is a large constant, with A positive when considering s positive and A negative in the other case. The solutions of (18) are of course $e^{\sqrt{A}s}$ and $e^{-\sqrt{A}s}$, the first growing rapidly, and the second decreasing rapidly as $s \to \infty$, if $A > 0$. A glance at the integration by parts following (16) shows that $\mathrm{Ai}(s)$ remains bounded when $s \to \infty$. So the comparison with $e^{\sqrt{A}s}$ must be dismissed, and we might reasonably guess that $\mathrm{Ai}(s)$ is rapidly decreasing in this case. When $s < 0$ we take $A < 0$ in (18). The exponentials $e^{\sqrt{A}s}$ and $e^{-\sqrt{A}s}$ are now oscillating, and we can therefore presume that $\mathrm{Ai}(s)$ should have an oscillatory character as $s \to -\infty$.

Theorem 3.1 *Suppose $u > 0$. Then as $u \to \infty$,*

(i) $\mathrm{Ai}(-u) = \pi^{-1/2}u^{-1/4}\cos(\frac{2}{3}u^{3/2} - \frac{\pi}{4})(1 + O(1/u^{3/4}))$.

(ii) $\mathrm{Ai}(u) = \dfrac{1}{2\pi^{1/2}}u^{-1/4}e^{-\frac{2}{3}u^{3/2}}(1 + O(1/u^{3/4}))$.

To consider the first case, we make the change of variables $x \mapsto u^{1/2}x$ in the defining integral with $s = -u$. This gives

$$\mathrm{Ai}(-u) = u^{1/2}I_-(u^{3/2}),$$

where

$$I_-(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it(x^3/3 - x)}\,dx. \tag{19}$$

Now write

$$I_-(s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-s\Phi(x)}\,dx,$$

where $\Phi(x) = \Phi_-(x) = x^3/3 - x$, and we shall apply Proposition 2.2, which in this case, since s is purely imaginary, is the method of stationary phase. Note that $\Phi'(x) = x^2 - 1$, so there are two critical points, $x_0 = \pm 1$; observe that $\Phi''(x) = 2x$; also $\Phi(\pm 1) = \mp 2/3$.

We break up the range of integration in (19) into two intervals $[-2, 0]$ and $[0, 2]$ each containing one critical point, and two complementary integrals, $(-\infty, -2]$ and $[2, \infty)$.

Now we apply Proposition 2.2 to the interval $[0, 2]$ with $s = -it$, $x_0 = 1$ $\psi = 1/2\pi$, $\Phi(1) = -2/3$, $\Phi''(1) = 2$, and get a contribution of

$$\frac{1}{2\sqrt{\pi}}e^{-i\frac{2}{3}t}\left(\frac{1}{(-it)^{1/2}} + O\left(\frac{1}{|t|}\right)\right),$$

in view of (8). Similarly the integral over $[-2, 0]$ contributes

$$\frac{1}{2\sqrt{\pi}} e^{i\frac{2}{3}t} \left(\frac{1}{(it)^{1/2}} + O\left(\frac{1}{|t|}\right) \right).$$

Finally, consider the complementary integrals. The first is

$$\int_{-\infty}^{-2} e^{it\Phi(x)}\, dx = \lim_{N\to\infty} \int_{-N}^{-2} e^{it\Phi(x)}\, dx = \lim_{N\to\infty} \frac{1}{it} \int_{-N}^{-2} \frac{d}{dx}\left(e^{it\Phi(x)}\right) \frac{dx}{\Phi'(x)},$$

where $\Phi'(x) = x^2 - 1$. So an integration by parts shows that this is $O(1/|t|)$. The integral over $[2, \infty)$ is treated similarly. Adding these four contributions, and inserting them in the identity $\mathrm{Ai}(-u) = u^{1/2} I_-(u^{3/2})$, proves conclusion (i) of the theorem.[6]

To deal with the conclusion (ii) of the theorem, we make the change of variables $x \mapsto u^{1/2}x$ in the integral (15), with $s = u$. This gives us, for $u > 0$,

$$\mathrm{Ai}(u) = u^{1/2} I_+(u^{3/2}),$$

where

$$I_+(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-sF(x)}\, dx \tag{20}$$

and $F(x) = -i(x^3/3 + x)$. Now when $s \to \infty$, the integrand in (20) again oscillates rapidly, but here in distinction to the previous case, there is no critical point on the real axis, since the derivative of $x^3/3 + x$ does not vanish. A repeated integration by parts argument (such as we have used before) shows that actually the integral $I_+(s)$ has fast decay as $s \to \infty$. But what is the exact nature and order of this decrease? To answer this question, we would have to take into account the precise cancellations inherent in (20), and doing this by the above method does not seem feasible.

A better way is to follow the guiding principle used in the asymptotics of the Bessel function, and to deform the line of integration in (20) to a contour on which the imaginary part of $F(z)$ vanishes; having done this, one might then hope to apply Laplace's method, Proposition 2.1, to find the true asymptotic behavior of $I_+(s)$, as $s \to \infty$.

We describe the idea in the more general situation in which we assume only that $F(z)$ is holomorphic. To follow the approach suggested, we seek a contour Γ so that:

---

[6] An alternative derivation of this conclusion can be given as a consequence of the relation of the Airy function with the Bessel functions. See Problem 3 below.

(a) $\mathrm{Im}(F) = 0$ on Γ.

(b) $\mathrm{Re}(F)$ has a minimum on Γ at some point z_0, and this function is non-degenerate in the sense that the second derivative of $\mathrm{Re}(F)$ along Γ is strictly positive at z_0.

Conditions (a) and (b) imply of course that $F'(z_0) = 0$. If as above, $F''(z_0) \neq 0$, then there are two curves Γ_1 and Γ_2 passing through z_0 which are orthogonal, so that $F|_{\Gamma_i}$ is real for $i = 1, 2$, with $\mathrm{Re}(F)$ restricted to Γ_1 having a minimum at z_0; and $\mathrm{Re}(F)$ restricted to Γ_2 having a maximum at z_0 (see Exercise 2 in Chapter 8). We therefore try to deform our original contour of integration to $\Gamma = \Gamma_1$. This approach is usually referred to as the method of **steepest descent**, because at z_0 the function $-\mathrm{Re}(F(z))$ has a saddle point, and starting at this point and following the path of Γ_1, one has the greatest decrease of this function.

Let us return to our special case, $F(z) = -i(z^3/3 + z)$. We note that

$$\begin{cases} \mathrm{Re}(F) &=& x^2y - y^3/3 + y, \\ \mathrm{Im}(F) &=& -x^3/3 + xy^2 - x. \end{cases}$$

We observe also that $F'(z) = -i(z^2 + 1)$, so we have two non-real critical points $z_0 = \pm i$ at which $F'(z_0) = 0$. If we choose $z_0 = i$, then the two curves passing through this point where $\mathrm{Im}(F) = 0$ are

$$\Gamma_1 = \{(x, y) : \; y^2 = x^2/3 + 1\} \quad \text{and} \quad \Gamma_2 = \{(x, y) : \; x = 0\}.$$

On Γ_2, the function $\mathrm{Re}(F)$ clearly has a maximum at the point $z_0 = i$, and so we reject this curve. We choose $\Gamma = \Gamma_1$, which is a branch of a hyperbola, and which can be written as $y = (x^2/3 + 1)^{1/2}$; it is asymptotic to the rays $z = re^{i\pi/6}$, and $z = re^{i5\pi/6}$ at infinity. See Figure 4.

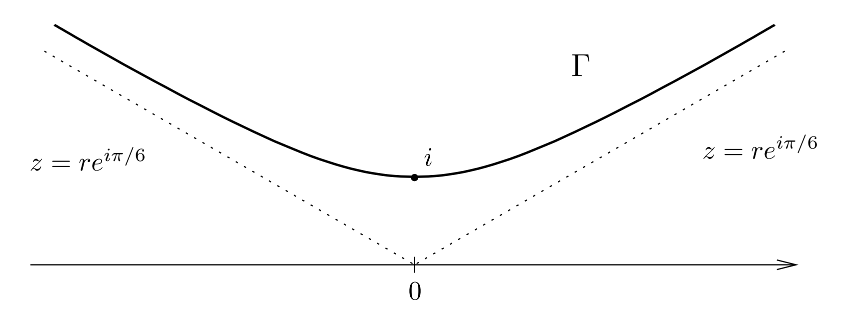


Figure 4. The curve of steepest descent

Next, we see that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-sF(x)}\,dx = \frac{1}{2\pi}\int_{\Gamma} e^{-sF(z)}\,dz. \tag{21}$$

This identity is justified by applying the Cauchy theorem to $e^{-sF(z)}$ on the contour Γ_R that consists of four arc segments: the parts of the real axis and Γ that lie inside the circle of radius R, and the two arcs of this circle joining the axis with Γ. Since in this region $e^{-sF(z)} = O(e^{-cyx^2})$ as $x \to \pm\infty$, the contributions of the two arcs of the circle are $O(\int_0^\pi e^{-cR^2\sin\theta}\,d\theta) = O(1/R)$, and letting $R \to \infty$ establishes (21).

We now observe that on Γ

$$\Phi(x) = \mathrm{Re}(F) = y(x^2 - y^2/3 + 1) = (\frac{8}{9}x^2 + \frac{2}{3})(x^2/3+1)^{1/2},$$

since $y^2 = x^2/3 + 1$ there. Also, on Γ we have that $dz = dx + idy = dx + i(x/3)(x^2/3+1)^{-1/2}dx$. Thus,

$$\frac{1}{2\pi}\int_{\Gamma} e^{-sF(z)}\,dz = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-s\Phi(x)}\,dx, \tag{22}$$

in view of the fact that $\Phi(x)$ is even, while $x(x^2/3+1)^{-1/2}$ is odd.

We note next that since $(1+u)^{1/2} = 1 + u/2 + O(u^2)$ as $u \to 0$,

$$\Phi(x) = (\frac{8}{9}x^2 + \frac{2}{3}) + \frac{2}{3}\frac{1}{2}\frac{x^2}{3} + O(x^4) = x^2 + \frac{2}{3} + O(x^4),$$

and so $\Phi''(0) = 2$. We now apply Proposition 2.1 to estimate the main part of the right-hand side of (22), by

$$\frac{1}{2\pi}\int_{-c}^{c} e^{-s\Phi(x)}\,dx,$$

where c is a small positive constant. Since $\Phi(0) = 2/3$, $\Phi''(0) = 2$, and $\psi(0) = 1/2\pi$, we obtain that this term contributes

$$e^{-\frac{2}{3}s}\left[\frac{1}{2\pi^{1/2}}\frac{1}{s^{1/2}} + O\left(\frac{1}{s}\right)\right].$$

The term $\int_c^\infty e^{-s\Phi(x)}\,dx$ is dominated by $e^{-2s/3}\int_c^\infty e^{-c_1 s x^2}\,dx$, which is $O(e^{-2s/3}e^{-\delta s})$ for some $\delta > 0$, as soon as $c > 0$. A similar estimate holds for $\int_{-\infty}^{-c} e^{-s\Phi(x)}\,dx$. Altogether, then,

$$I_+(s) = e^{-\frac{2}{3}s}\left[\frac{1}{2\pi^{1/2}}\frac{1}{s^{1/2}} + O\left(\frac{1}{s}\right)\right] \qquad \text{as } s \to \infty,$$

and this gives the desired asymptotic (ii) for the Airy function.

4 The partition function

Our last illustration of the techniques developed in this appendix is in their application to the partition function $p(n)$, which was discussed in Chapter 10. We derive for it the main term of the remarkable asymptotic formula of Hardy-Ramanujan.

Theorem 4.1 *If p denotes the partition function, then*

(i) $p(n) \sim \dfrac{1}{4\sqrt{3}n} e^{Kn^{1/2}}$ *as $n \to \infty$, where $K = \pi\sqrt{\frac{2}{3}}$.*

(ii) *A much more precise assertion is that*

$$p(n) = \frac{1}{2\pi\sqrt{2}} \frac{d}{dn}\left(\frac{e^{K(n-\frac{1}{24})^{1/2}}}{(n-\frac{1}{24})^{1/2}}\right) + O(e^{\frac{K}{2}n^{1/2}}).$$

Note. Observe that $(n - \frac{1}{24})^{1/2} - n^{1/2} = O(n^{-1/2})$, by the mean-value theorem; hence $e^{K(n-\frac{1}{24})^{1/2}} = e^{Kn^{1/2}}(1 + O(n^{-1/2}))$, thus $e^{K(n-\frac{1}{24})^{1/2}} \sim e^{Kn^{1/2}}$, as $n \to \infty$. Of course, clearly $(n - \frac{1}{24})^{1/2} \sim n^{1/2}$, and in particular (ii) implies (i).

We shall discuss first, in a more general setting, how we might derive the asymptotic behavior of a sequence $\{F_n\}$ from the analytic properties of its generating function $F(w) = \sum_{n=0}^{\infty} F_n w^n$. Assuming for the sake of simplicity that $\sum F_n w^n$ has the unit disc as its disc of convergence, we can set forth the following heuristic principle: the asymptotic behavior of F_n is determined by the location and nature of the "singularities" of F on the unit circle, and the contribution to the asymptotic formula due to each singularity corresponds in magnitude to the "order" of that singularity.

A very simple example in which this principle is unambiguous and can be verified occurs when F is meromorphic in a larger disc, but has only one singularity on the circle, a pole of order r at the point $w = 1$. Then there is a polynomial P of degree $r - 1$ so that $F_n = P(n) + O(e^{-\epsilon n})$ as $n \to \infty$, for some $\epsilon > 0$. In fact, $\sum_{n=0}^{\infty} P(n)w^n$ is a good approximation to $F(w)$ near $w = 1$; it is the principal part of the pole of F. (See also Problem 4.)

For the partition function the analysis is not as simple as this example, but the principle stated above is still applicable when properly interpreted. To this task we now turn.

We recall the formula

$$\sum_{n=0}^{\infty} p(n)\, w^n = \prod_{n=1}^{\infty} \frac{1}{1 - w^n},$$

established in Theorem 2.1, Chapter 10. This identity implies that the generating function is holomorphic in the unit disc. In what follows, it will be convenient to pass from the unit disc to the upper half-plane by writing $w = e^{2\pi iz}$, $z = x + iy$, and taking $y > 0$. We therefore have

$$\sum_{n=0}^{\infty} p(n)e^{2\pi inz} = f(z),$$

with

$$f(z) = \prod_{n=1}^{\infty} \frac{1}{1 - e^{2\pi inz}},$$

and

$$p(n) = \int_{\gamma} f(z)e^{-2\pi inz}\, dz. \tag{23}$$

Here γ is the segment in the upper half-plane joining $-1/2 + i\delta$ to $1/2 + i\delta$, with $\delta > 0$; the height δ will be fixed later in terms of n.

To proceed further, we look first at where the main contribution to the integral (23) might be, in terms of the relative size of $f(x + iy)$, as $y \to 0$. Notice that f is largest near $z = 0$. This is because $|f(x + iy)| \le f(iy)$, and moreover $f(iy)$ increases as y decreases, in view of the fact that the coefficients $p(n)$ are positive. Alternatively, we observe that each factor $1 - e^{2\pi inz}$, appearing in the product for f, vanishes as $z \to 0$, but the same is true for any other point (mod 1) on the real axis. Thus in analogy with the simple example considered above, we seek an elementary function f_1, which has much the same behavior as f at $z = 0$, and try to replace f by f_1 in (23).

It is here that we are very fortunate, because the generating function is just a variant of the Dedekind eta function,

$$\eta(z) = e^{i\pi z/12} \prod_{n=1}^{\infty} (1 - e^{2\pi inz}).$$

From this, it is obvious that

$$f(z) = e^{\frac{i\pi z}{12}} (\eta(z))^{-1}.$$

(Incidentally, the fraction $1/12$ arising above will explain the occurrence of the fraction $1/24$ in the asymptotic formula for $p(n)$.)

Since η satisfies the functional equation $\eta(-1/z) = \sqrt{z/i}\ \eta(z)$ (see Proposition 1.9 in Chapter 10), it follows that

$$f(z) = \sqrt{z/i}\ e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}} f(-1/z). \tag{24}$$

Notice also that if z is appropriately restricted and $z \to 0$, then $\mathrm{Im}(-1/z) \to \infty$, from which it follows that $f(-1/z) \to 1$ rapidly, because

$$f(z) = 1 + O(e^{-2\pi y}), \qquad z = x + iy,\ y \geq 1. \tag{25}$$

Thus it is natural to choose $f_1(z) = \sqrt{z/i}\ e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}}$ as the function which approximates well the generating function $f(z)$ (at $z = 0$), and write (because of (24))

$$p(n) = p_1(n) + E(n),$$

with

$$\begin{cases} p_1(n) &= \displaystyle\int_\gamma \sqrt{z/i}\ e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}} e^{-2\pi i n z}\, dz, \\ E(n) &= \displaystyle\int_\gamma \sqrt{z/i}\ e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}} e^{-2\pi i n z} (f(-1/z) - 1)\, dz. \end{cases}$$

We first take care of the error term $E(n)$, and in doing so we specify γ by choosing its height in terms of n. In estimating $E(n)$ we replace its integrand by its absolute value and note that if $z \in \gamma$, then

$$\left| \sqrt{z/i}\ e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}} e^{-2\pi i n z} \right| \leq c e^{2\pi n \delta} e^{\frac{\pi}{12} \frac{\delta}{\delta^2 + x^2}}, \tag{26}$$

since $z = x + iy$, and $\mathrm{Re}(i/z) = \delta/(\delta^2 + x^2)$.

On the other hand, we can make two estimates for $f(-1/z) - 1$. The first arises from (25) by replacing z by $-1/z$, and gives

$$|f(-1/z) - 1| \leq c e^{-2\pi \frac{\delta}{\delta^2 + x^2}} \qquad \text{if } \tfrac{\delta}{\delta^2 + x^2} \geq 1. \tag{27}$$

For the second, we observe that $|f(z)| \leq f(iy) \leq C e^{\frac{\pi}{12y}}$, when $y \leq 1$, because of the functional equation (24), and hence

$$|f(-1/z) - 1| \leq O\left(e^{\frac{\pi}{12} \frac{\delta^2 + x^2}{\delta}}\right) = O\left(e^{\frac{\pi}{48\delta}}\right) \tag{28}$$

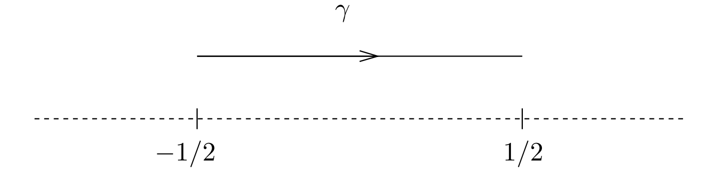
if $\frac{\delta}{\delta^2 + x^2} \leq 1$, since $|x| \leq 1/2$.

Therefore in the integral defining $E(n)$ we use (26) and (27) when $\frac{\delta}{\delta^2+x^2} \geq 1$, and (26) and (28) when $\frac{\delta}{\delta^2+x^2} \leq 1$. The first leads to a contribution of $O(e^{2\pi n\delta})$, since $2\pi > \pi/12$. The second gives a contribution of $O(e^{2\pi n\delta} e^{\frac{\pi}{48\delta}})$. Hence $E(n) = O(e^{2\pi n\delta} e^{\frac{\pi}{48\delta}})$, and we choose δ so as to minimize the right-hand side, that is, $2\pi n\delta = \frac{\pi}{48\delta}$; this means we take $\delta = \frac{1}{4\sqrt{6}\, n^{1/2}}$, and we get

$$E(n) = O\left(e^{\frac{4\pi}{4\sqrt{6}}\, n^{1/2}}\right) = O\left(e^{\frac{K}{2}\, n^{1/2}}\right),$$

which is the desired size of the error term.

We turn to the main term $p_1(n)$. To simplify later calculations we "improve" the contour γ by adding to it two small end-segments; these are the segment joining $-1/2$ to $-1/2 + i\delta$ and that joining $1/2 + i\delta$ to $1/2$. We call this new contour γ' (see Figure 5).



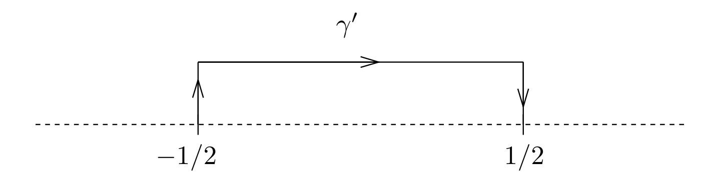


Figure 5. γ and the improved contour γ'

Notice that since $\sqrt{z/i}\, e^{\frac{i\pi}{12z}}$ is $O(1)$ on the two added segments (for the integral defining $p_1(n)$), the modification contributes $O(e^{2\pi n\delta}) = O(e^{\frac{2\pi}{4\sqrt{6}} n^{1/2}}) = O(e^{\frac{K}{4} n^{1/2}})$, which is even smaller than the allowed error, and therefore can be incorporated in $E(n)$. So without introducing further notation we will rewrite $p_1(n)$ replacing the contour γ by γ' in the integration defining p_1, namely

$$p_1(n) = \int_{\gamma'} \sqrt{z/i}\; e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}} e^{-2\pi i n z}\, dz. \tag{29}$$

Next we simplify the triad of exponentials appearing in (29) by making a change of variables $z \mapsto \mu z$ so that their combination takes the form

$$e^{Ai\left(\frac{1}{z}-z\right)}.$$

This can be achieved under the two conditions $A = 2\pi\mu(n - \frac{1}{24})$ and $A = \frac{\pi}{12\mu}$, which means that

$$A = \frac{\pi}{\sqrt{6}}\,(n - \frac{1}{24})^{1/2} \quad \text{and} \quad \mu = \frac{1}{2\sqrt{6}}\,(n - \frac{1}{24})^{-1/2}.$$

Making the indicated change of variables we now have

$$p_1(n) = \mu^{3/2} \int_\Gamma e^{-sF(z)}\ \sqrt{z/i}\ dz, \tag{30}$$

with $F(z) = i(z - 1/z)$, $s = \frac{\pi}{\sqrt{6}}\,(n - \frac{1}{24})^{1/2}$. The curve Γ (see Figure 6). is now the union of three segments $[-a_n, -a_n + i\delta']$, $[-a_n + i\delta', a_n + i\delta']$, and $[a_n + i\delta', a_n]$; we can write $\Gamma = \mu^{-1}\gamma'$.

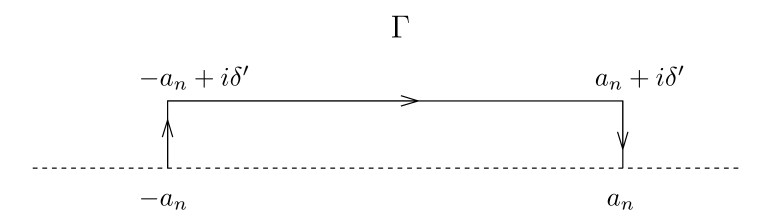


Figure 6. The curve Γ

Here $a_n = \frac{1}{2}\mu^{-1} = \sqrt{6}\,(n - \frac{1}{24})^{1/2} \approx n^{1/2}$, while $\delta' = \delta\mu^{-1} = \frac{2\sqrt{6}}{4\sqrt{6}n^{1/2}}\,(n - \frac{1}{24})^{1/2} \sim 1/2$, as $n \to \infty$.

We apply the method of steepest descent to the integral (30). In doing this, we note that $F(z) = i(z - 1/z)$ has one (complex) critical point $z = i$, in the upper half-plane. Moreover, the two curves passing through i on which F is real are: the imaginary axis, on which F has a maximum at $z = i$, which we reject, and the unit circle, on which F has a minimum at $z = i$. Thus using Cauchy's theorem we replace the integration on Γ by the integration over our final curve Γ^*, which consists of the segment $[-a_n, -1]$, $[1, a_n]$, together with the upper semicircle joining -1 to 1.

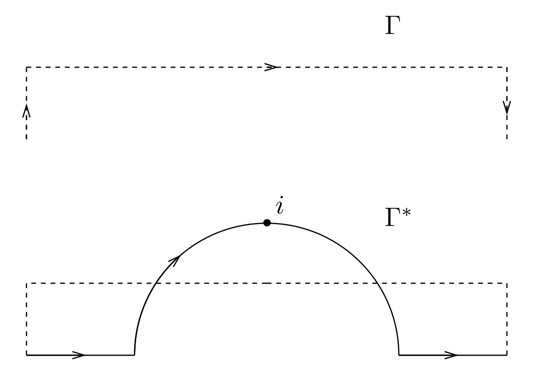


Figure 7. The final curve Γ^*

We therefore have

$$p_1(n) = \mu^{3/2} \int_{\Gamma^*} e^{-sF(z)} \sqrt{z/i}\, dz.$$

The contributions on the segments $[-a_n, -1]$ and $[1, a_n]$ are relatively very small, because on the real axis the exponential has absolute value 1, and hence the integrand is bounded by $\sup_{|z| \le a_n} |z|^{1/2}$, and this leads to two terms which are $O(a_n^{3/2} \mu^{3/2}) = O(1)$.

Finally, we come to the principal part, which is the integration over the semicircle, taken with the orientation on the figure. Here we write $z = e^{i\theta}$, $dz = ie^{i\theta}\, d\theta$. Since $i(z - 1/z) = -2\sin\theta$, this gives a contribution

$$-\mu^{3/2}\int_0^{\pi} e^{2s\sin\theta} e^{i3\theta/2}\sqrt{i}\, d\theta = \mu^{3/2}\int_{-\pi/2}^{\pi/2} e^{2s\cos\theta}(\cos(3\theta/2) + i\sin(3\theta/2))d\theta.$$

In applying Proposition 2.1, Laplace's method, we take $\Phi(\theta) = -\cos\theta$, $\theta_0 = 0$, so $\Phi(\theta_0) = -1$, $\Phi''(\theta_0) = 1$ and we choose $\psi(\theta) = \cos(3\theta/2) + i\sin(3\theta/2)$, so that $\psi(\theta_0) = 1$. Therefore, the above contributes

$$\mu^{3/2} e^{2s} \frac{\sqrt{2\pi}}{(2s)^{1/2}} \left(1 + O(s^{-1/2})\right).$$

Now since $s = \frac{\pi}{\sqrt{6}}(n - \frac{1}{24})^{1/2}$, $\frac{2\pi}{\sqrt{6}} = \pi\sqrt{\frac{2}{3}} = K$, and $\mu = \frac{\sqrt{6}}{12}(n - \frac{1}{24})^{-1/2}$, we obtain

$$p(n) = \frac{1}{4n\sqrt{3}}\, e^{K\, n^{1/2}} \left(1 + O\left(n^{-1/4}\right)\right),$$

and the first conclusion of the theorem is established.

To obtain the more exact conclusion (ii), we retrace our steps and use an additional device, which allows us to evaluate rather precisely the key integral. With $p_1(n)$ defined by (29), which is an integral taken over $\gamma' = \gamma'_n$, we write

$$p_1(n) = \frac{d}{dn}q(n) + e(n),$$

where

$$q(n) = \frac{1}{2\pi}\int_{\gamma'} (z/i)^{-1/2} e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}} e^{-2\pi i n z}\, dz,$$

and $e(n)$ is the term due to the variation of the contour $\gamma' = \gamma'_n$, when forming the derivative in n. By Cauchy's theorem this is easily seen to be dominated by $O(e^{2\pi n\delta})$, which we have seen is $O(e^{\frac{K}{4}n^{1/2}})$, and can be subsumed in the error term. To analyze $q(n)$, we proceed as before, first making the change of variables $z \mapsto \mu z$, and then replacing the resulting contour Γ by Γ^*. As a consequence, we have

$$q(n) = \frac{\mu^{1/2}}{2\pi}\int_{\Gamma^*} e^{-sF(z)}(z/i)^{-1/2}\, dz, \tag{31}$$

with $F(z) = i(z - 1/z)$, $s = \frac{\pi}{\sqrt{6}}(n - \frac{1}{24})^{1/2}$, and $\mu = \frac{1}{2\sqrt{6}}(n - \frac{1}{24})^{-1/2}$.

Now the two segments $[-a_n, -1]$ and $[1, a_n]$ of the contour Γ^* make harmless contributions to $\frac{d}{dn}q(n)$, since F is purely imaginary on the real axis. Indeed, they yield terms which are $O(a_n^{1/2}\mu^{1/2}) = O(1)$.

The main part of (31) is the term arising from the integration on the semicircle. Thus setting $z = e^{i\theta}$, $dz = ie^{i\theta}\, d\theta$, and $i(z - 1/z) = -2\sin\theta$, it equals

$$\begin{aligned} -\frac{\mu^{1/2}}{2\pi}\int_0^{\pi} e^{2s\sin\theta} e^{i\theta/2} i^{3/2}\, d\theta &= \frac{\mu^{1/2}}{2\pi}\int_{-\pi/2}^{\pi/2} e^{2s\cos\theta}(\cos(\theta/2) + i\sin(\theta/2))d\theta \\ &= \frac{\mu^{1/2}}{2\pi}\int_{-\pi/2}^{\pi/2} e^{2s\cos\theta}\cos(\theta/2)d\theta, \end{aligned}$$

where we have used the fact that the integral $\int_{-\pi/2}^{\pi/2} e^{2s\cos\theta}\sin(\theta/2)\, d\theta$ vanishes since the integrand is odd.

Now $\cos\theta = 1 - 2(\sin\theta/2)^2$, so setting $x = \sin(\theta/2)$ we see that the above integral becomes

$$\frac{\mu^{1/2}e^{2s}}{\pi}\int_{-\frac{\sqrt{2}}{2}}^{\frac{\sqrt{2}}{2}} e^{-4sx^2}\,dx.$$

However

$$\begin{aligned}\int_{-\frac{\sqrt{2}}{2}}^{\frac{\sqrt{2}}{2}} e^{-4sx^2}\,dx &= \int_{-\infty}^{\infty} e^{-4sx^2}\,dx + O\left(\int_{\frac{\sqrt{2}}{2}}^{\infty} e^{-4sx^2}\,dx\right)\\ &= \frac{\sqrt{\pi}}{2s^{1/2}} + O(e^{-2s}),\end{aligned}$$

and also

$$\frac{d}{ds}\left(\int_{-\frac{\sqrt{2}}{2}}^{\frac{\sqrt{2}}{2}} e^{-4sx^2}\,dx\right) = \frac{d}{ds}\left(\frac{\sqrt{\pi}}{2s^{1/2}}\right) + O(e^{-2s}).$$

Gathering all the error terms together, we find

$$p(n) = \frac{d}{dn}\left(\mu^{1/2}\frac{e^{2s}}{\pi}\frac{\sqrt{\pi}}{2s^{1/2}}\right) + O(e^{\frac{K}{2}n^{1/2}}).$$

Since $s = \frac{\pi}{\sqrt{6}}(n-\frac{1}{24})^{1/2}$, $\mu = \frac{\sqrt{6}}{12}(n-\frac{1}{24})^{-1/2}$, and $K = \pi\sqrt{\frac{2}{3}}$, this is

$$p(n) = \frac{1}{2\pi\sqrt{2}}\frac{d}{dn}\left(\frac{e^{K(n-\frac{1}{24})^{1/2}}}{(n-\frac{1}{24})^{1/2}}\right) + O(e^{\frac{K}{2}n^{1/2}}),$$

and the theorem is proved.

5 Problems

1. Let η be an indefinitely differentiable function supported in a finite interval, so that $\eta(x) = 1$ for x near 0. Then, if $m > -1$ and $N > 0$,

$$\int_0^\infty e^{-sx}x^m\eta(x)\,dx = s^{-m-1}\Gamma(m+1) + O(s^{-N})$$

for $\mathrm{Re}(s) \geq 0$, $|s| \to \infty$.

(a) Consider first the case $-1 < m \leq 0$. It suffices to see that

$$\int_0^\infty e^{-sx}x^m(1-\eta(x))\,dx = O(s^{-N}),$$

and this can be done by repeated integration by parts since $e^{-sx} = (-1)^N s^{-N}\left(\frac{d}{dx}\right)^N(e^{-sx})$.

(b) To extend this to all m, find an integer k so that $k-1<m\leq k$, write

$$\int\left[\left(\frac{d}{dx}\right)^k(x^m)\right]e^{-sx}\eta(x)\,dx=c_{k,m}s^{-m+k-1}+O(s^{-N}),$$

and integrate by parts k times.

2. The following is a more precise version of Stirling's formula. There are real constant $a_1=1/12$, a_2, ..., a_n, ..., so that for every $N>0$

$$\Gamma(s)=e^{s\log s}e^{-s}\frac{\sqrt{2\pi}}{s^{1/2}}\left(1+\sum_{j=1}^{N}a_js^{-j}+O(s^{-N})\right)\qquad\text{when } s\in S_\delta.$$

This can be proved by using the results of Problem 1 in place of Proposition 1.3.

3. The Bessel functions and Airy function have the following power series expansions:

$$J_\nu(x)=\left(\frac{x}{2}\right)^\nu\sum_{m=0}^{\infty}\frac{(-1)^m\left(\frac{x^2}{4}\right)^m}{m!\Gamma(\nu+m+1)},$$

$$\mathrm{Ai}(-x)=\frac{1}{\pi}\sum_{n=0}^{\infty}\frac{x^n}{n!}\sin(2\pi(n+1)/3)3^{n/3-2/3}\Gamma(n/3+1/3).$$

(a) From this, verify that when $x>0$,

$$\mathrm{Ai}(-x)=\frac{x^{1/2}}{3}\left(J_{1/3}\left(\frac{2}{3}x^{3/2}\right)+J_{-1/3}\left(\frac{2}{3}x^{3/2}\right)\right).$$

(b) The function $\mathrm{Ai}(x)$ extends to an entire function of order $3/2$.

[Hint: For (b), use (a), or alternatively, apply Problem 4 in Chapter 5 to the power series for Ai. Compare also with Problem 1, Chapter 4.]

4. Suppose $F(z)=\sum_{n=0}^{\infty}F_nw^n$ is meromorphic in a region containing the closed unit disc, and the only poles of F are on the unit circle at the points $\alpha_1,\ldots,\alpha_k$, and their orders are $r_1,\ldots,r_k$ respectively. Then for some $\epsilon>0$

$$F_n=\sum_{j=1}^{k}P_j(n)+O(e^{-\epsilon n})\qquad\text{as } n\to\infty.$$

Here

$$P_j(n)=\frac{1}{(r_j-1)!}\left(\frac{d}{dw}\right)^{r_j-1}\left[(w-\alpha_j)^{r_j}w^{-n-1}F(w)\right]_{w=\alpha_j}.$$

Note that each P_j is of the form $P_j(n) = A_j(\alpha_j^{-1}n)^{r_j-1} + O(n^{r_j-2})$.

To prove this, use the residue formula (Theorem 1.4, Chapter 3).

5.* The one shortcoming in our derivation of the asymptotic formula for $p(n)$ arose from the fact that while $f_1(z) = \sqrt{z/i}\ e^{\frac{i\pi}{12z}} e^{\frac{i\pi z}{12}}$ is a good approximation to the generating function $f(z)$ near $z = 0$, this fails near other points on the real axis, since f_1 is regular there, but f is not.

However, using the transformation law (24) and the identity $f(z+1) = f(z)$, one can derive the following generalization of (24): whenever p/q is a rational number in lowest form (so p and q are relatively prime) then

$$f\left(z - \frac{p}{q}\right) = \omega_{p/q}\sqrt{\frac{zq}{i}}\ e^{\frac{i\pi}{12zq^2}} e^{-\frac{i\pi z}{12}} f\left(-\frac{1}{zq^2} - \frac{p'}{q}\right),$$

where $pp' = 1 \mod q$. Here $\omega_{p/q}$ is an appropriate 24th root of unity. This formula leads to an analogous $f_{p/q}$, approximating f at $z = p/q$.

From this one can obtain for each p/q a contribution of the form

$$c_{p/q}\frac{1}{2\pi\sqrt{2}}\frac{d}{dn}\left(\frac{e^{\frac{K}{q}(n-\frac{1}{24})^{1/2}}}{(n-\frac{1}{24})^{1/2}}\right)$$

to the asymptotic formula for $p(n)$. When suitably modified, the resulting series, summed over all proper fractions p/q in $[0,1)$, actually converges and gives an exact formula for $p(n)$.

Appendix B: Simple Connectivity and Jordan Curve Theorem

> Jordan was one of the precursors of the theory of functions of a real variable. He introduced in this part of analysis the capital notion of functions of bounded variation. Not less celebrated is his study of curves, universally called Jordan curves, which curves separate the plane in two distinct regions. We also owe him important propositions regarding the measure of sets that have led the way to numerous modern researches.
>
> *E. Picard,* 1922

The notion of simple connectivity is at the source of many basic and fundamental results in complex analysis. To clarify the meaning of this important concept, we have gathered in this appendix some further insights into the properties of simply connected sets. Closely tied to the idea of simple connectivity is the notion of the "interior" of a simple closed curve. The theorem of Jordan states that this interior is well-defined and is simply connected. We prove here the special case of this theorem for curves which are piecewise-smooth.

Recall the definition in Chapter 3, according to which a region Ω is simply connected if any two curves in Ω with the same end-points are homotopic. From this definition we deduced an important version of Cauchy's theorem which states that if Ω is simply connected and $\gamma \subset \Omega$ is any closed curve, then

$$\int_\gamma f(\zeta)\, d\zeta = 0 \tag{1}$$

whenever f is holomorphic in Ω. Here, we shall prove that a converse also holds, therefore:

(I) A region Ω is simply connected if and only if it is holomorphically simply connected; that is, whenever $\gamma \subset \Omega$ is closed and f holomorphic in Ω then (1) holds.

Besides this fundamental equivalence, which is analytic in nature, there

are also topological conditions that can be used to describe simple connectivity. In fact, the definition in terms of homotopies suggests that a simply connected set has no "holes." In other words, one cannot find a closed curve in Ω that loops around points that do not belong to Ω. In the first part of this appendix we shall also turn these intuitive statements into tangible theorems:

(II) We show that a bounded region Ω is simply connected if and only if its complement is connected.

(III) We define the winding number of a curve around a point, and prove that Ω is simply connected if and only if no curve in Ω winds around points in the complement of Ω.

In the second part of this appendix we return to the problem of curves and their interior. The main question is the following: given a closed curve Γ that does not intersect itself (it is simple), can we make sense of the "region enclosed by Γ"? In other words, what is the "interior" of Γ? Naturally, we may expect the interior to be open, bounded, simply connected, and have Γ as its boundary. To solve this problem, at least when the curve is piecewise-smooth, we prove a theorem that guarantees the existence of a unique set which satisfies all the desired properties. This is a special case of the Jordan curve theorem, which is valid in the general case when the simple curve is assumed to be merely continuous. In particular, our result leads to a generalization of Cauchy's theorem in Chapter 2 which we formulated for toy contours.

We continue to follow the convention set in Chapter 1 by using the term "curve" synonymously with "piecewise-smooth curve," unless stated otherwise.

1 Equivalent formulations of simple connectivity

We first dispose of (I).

Theorem 1.1 *A region Ω is holomorphically simply connected if and only if Ω is simply connected.*

Proof. One direction is simply the version of Cauchy's theorem in Corollary 5.3, Chapter 3. Conversely, suppose that Ω is holomorphically simply connected. If $\Omega = \mathbb{C}$, then it is clearly simply connected. If Ω is not all of $\mathbb{C}$, recall that the Riemann mapping theorem still applies (see the remark following its proof in Chapter 8), hence Ω is conformally equivalent to the unit disc. Since the unit disc is simply connected, the same must be true of Ω.

Next, we turn to (II) and (III), which, as we mentioned, are both precise formulations of the fact that a simply connected region cannot have "holes."

Theorem 1.2 *If Ω is a bounded region in $\mathbb{C}$, then Ω is simply connected if and only if the complement of Ω is connected.*

Note that we assume that Ω is bounded. If this is not the case, then the theorem as stated does not hold, for example an infinite strip is simply connected yet its complement consists of two components. However, if the complement is taken with respect to the extended complex plane, that is, the Riemann sphere, then the conclusion of the theorem holds regardless of whether Ω is bounded or not.

Proof. We begin with the proof that if Ω^c is connected, then Ω is simply connected. This will be achieved by showing that Ω is holomorphically simply connected. Therefore, let γ be a closed curve in Ω and f a holomorphic function on Ω. Since Ω is bounded, the set[1]

$$K = \{z \in \Omega : d(z, \Omega^c) \geq \epsilon\}$$

is compact, and for sufficiently small ϵ, the set K contains γ. In an attempt to apply Runge's theorem (Theorem 5.7 in Chapter 2), we must first show that the complement K^c of K is connected.

If this is not the case, then K^c is the disjoint union of two non-empty open sets, say $K^c = \mathcal{O}_1 \cup \mathcal{O}_2$. Let

$$F_1 = \mathcal{O}_1 \cap \Omega^c \quad \text{and} \quad F_2 = \mathcal{O}_2 \cap \Omega^c.$$

Clearly, $\Omega^c = F_1 \cup F_2$, so if we can show that F_1 and F_2 are disjoint, closed, and non-empty, then we will conclude that Ω^c is not connected, thus contradicting the hypothesis in the theorem. Since $\mathcal{O}_1$ and $\mathcal{O}_2$ are disjoint, so are F_1 and F_2. To see why F_1 is closed, suppose $\{z_n\}$ is a sequence of points in F_1 that converges to z. Since Ω^c is closed we must have $z \in \Omega^c$, and since Ω^c is at a finite distance from K, we deduce that $z \in \mathcal{O}_1 \cup \mathcal{O}_2$. Now we observe that we cannot have $z \in \mathcal{O}_2$, for otherwise we would have $z_n \in \mathcal{O}_2$ for sufficiently large n because $\mathcal{O}_2$ is open, and this contradicts the fact that $z_n \in F_1$ and $\mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$. Hence $z \in \mathcal{O}_1$ and F_1 is closed, as desired. Finally, we claim that F_1 is non-empty. If otherwise, $\mathcal{O}_1$ is contained in Ω. Select any point $w \in \mathcal{O}_1$, and since $w \notin K$, there exists $z \in \Omega^c$ with $|w - z| < \epsilon$, and the entire line segment from w to z belongs to K^c. Since $z \in \mathcal{O}_2$ (because $\mathcal{O}_1 \subset \Omega$), some point

[1] Here, $d(z, \Omega^c) = \inf\{|z - w| : \, w \in \Omega^c\}$ denotes the distance from z to Ω^c.

on the line segment $[z, w]$ must belong to neither $\mathcal{O}_1$ nor $\mathcal{O}_2$, and this is a contradiction. More precisely, if we set

$$t^* = \sup\{0 \leq t \leq 1 : (1-t)z + tw \in \mathcal{O}_2\},$$

then $0 < t^* < 1$, and the point $(1-t^*)z + t^*w$, which is not in K, cannot belong to either $\mathcal{O}_1$ or $\mathcal{O}_2$ since these sets are open. Similar arguments imply the same conclusions for F_2, and we have reached the desired contradiction. Thus K^c is connected.

Therefore, Runge's theorem guarantees that f can be approximated uniformly on K, and hence on γ, by polynomials. However, $\int_\gamma P(z)\,dz = 0$ whenever P is a polynomial, so in the limit we conclude that $\int_\gamma f(z)\,dz = 0$, as desired.

The converse result, that Ω^c is connected whenever Ω is bounded and simply connected, will follow from the notion of winding numbers, which we discuss next.

Winding numbers

If γ is a closed curve in $\mathbb{C}$ and z a point not lying on γ, then we may calculate the number of times the curve γ winds around z by looking at the change of argument of the quantity $\zeta - z$ as ζ travels on γ. Every time γ loops around z, the quantity $(1/2\pi)\arg(\zeta - z)$ increases (or decreases) by 1. If we recall that $\log w = \log|w| + i \arg w$, and denote the beginning and ending points of γ by ζ_1 and ζ_2, then we may guess that the quantity

$$\frac{1}{2\pi i}\left[\log(\zeta_1 - z) - \log(\zeta_2 - z)\right], \quad \text{which “equals”} \quad \frac{1}{2\pi i}\int_\gamma \frac{d\zeta}{\zeta - z},$$

computes precisely the number of times γ loops around ζ.

These considerations lead to the following precise definition: the **winding number** of a closed curve γ around a point $z \notin \gamma$ is

$$W_\gamma(z) = \frac{1}{2\pi i}\int_\gamma \frac{d\zeta}{\zeta - z}.$$

Sometimes, $W_\gamma(z)$ is also called the index of z with respect to γ.

For example, if $\gamma(t) = e^{ikt}$, $0 \leq t \leq 2\pi$, is the unit circle traversed k times in the positive direction (with $k \in \mathbb{N}$), then $W_\gamma(0) = k$. In fact, one has

$$W_\gamma(z) = \begin{cases} k & \text{if } |z| < 1, \\ 0 & \text{if } |z| > 1. \end{cases}$$

Similarly, if $\gamma(t) = e^{-ikt}$, $0 \le t \le 2\pi$, is the unit circle traversed k times in the negative direction, then we find that $W_\gamma(z) = -k$ in the interior of the disc, and $W_\gamma(z) = 0$ in its exterior.

Note that, if γ denotes a positively oriented toy contour, then

$$W_\gamma(z) = \begin{cases} 1 & \text{if } z \in \text{interior of } \gamma, \\ 0 & \text{if } z \in \text{exterior of } \gamma. \end{cases}$$

In general we have the following natural facts about winding numbers.

Lemma 1.3 *Let γ be a closed curve in $\mathbb{C}$.*

(i) *If $z \notin \gamma$, then $W_\gamma(z) \in \mathbb{Z}$.*

(ii) *If z and w belong to the same open connected component in the complement of γ, then $W_\gamma(z) = W_\gamma(w)$.*

(iii) *If z belongs to the unbounded connected component in the complement of γ, then $W_\gamma(z) = 0$.*

Proof. To see why (i) is true, suppose that $\gamma : [0,1] \to \mathbb{C}$ is a parametrization for the curve, and let

$$G(t) = \int_0^t \frac{\gamma'(s)}{\gamma(s) - z}\, ds.$$

Then G is continuous and, except possibly at finitely many points, it is differentiable with $G'(t) = \gamma'(t)/(\gamma(t) - z)$. This implies that, except possibly at finitely many points, the derivative of the continuous function $H(t) = (\gamma(t) - z)e^{-G(t)}$ is zero, and hence H must be constant. Putting $t = 0$ and recalling that γ is closed, so that $\gamma(0) = \gamma(1)$, we find

$$1 = e^{G(0)} = c(\gamma(0) - z) = c(\gamma(1) - z) = e^{G(1)}.$$

Therefore, $G(1)$ is an integral multiple of $2\pi i$, as desired.

For (ii), we simply note that $W_\gamma(z)$ is a continuous function of $z \notin \gamma$ that is integer-valued, so it must be constant in any open connected component in the complement of γ.

Finally, one observes that $\lim_{|z|\to\infty} W_\gamma(z) = 0$, and, combined with (ii), this establishes (iii).

We now show that the notion of a bounded simply connected set Ω may be understood in the following sense: no curve in Ω winds around points in Ω^c.

Theorem 1.4 *A bounded region Ω is simply connected if and only if $W_\gamma(z) = 0$ for any closed curve γ in Ω and any point z not in Ω.*

Proof. If Ω is simply connected and $z \notin \Omega$, then $f(\zeta) = 1/(\zeta - z)$ is holomorphic in Ω, and Cauchy's theorem gives $W_\gamma(z) = 0$.

For the converse, it suffices to prove that the complement of Ω is connected (Theorem 1.2). We argue by contradiction, and construct an explicit closed curve γ in Ω and find a point w so that $W_\gamma(w) \neq 0$.

If we suppose that Ω^c is not connected, then we may write $\Omega^c = F_1 \cup F_2$ where F_1, F_2 are disjoint, closed, and non-empty. Only one of these sets can be unbounded, so that we may assume that F_1 is bounded, thus compact. The curve γ will be constructed as part of the boundary of an appropriate union of squares.

Lemma 1.5 *Let w be any point in F_1. Under the above assumptions, there exists a finite collection of closed squares $\mathcal{Q} = \{Q_1, \ldots, Q_n\}$ that belong to a uniform grid $\mathcal{G}$ of the plane, and are such that:*

(i) *w belongs to the interior of Q_1.*

(ii) *The interiors of Q_j and Q_k are disjoint when $j \neq k$.*

(iii) *F_1 is contained in the interior of $\bigcup_{j=1}^n Q_j$.*

(iv) *$\bigcup_{j=1}^n Q_j$ is disjoint from F_2.*

(v) *The boundary of $\bigcup_{j=1}^n Q_j$ lies entirely in Ω, and consists of a finite union of disjoint simple closed polygonal curves.*

Assuming this lemma for now, we may easily finish the proof of the theorem. The boundary ∂Q_j of each square is equipped with the positive orientation. Since $w \in Q_1$, and $w \notin Q_j$ for all $j > 1$, we have

$$\sum_{j=1}^n \frac{1}{2\pi i} \int_{\partial Q_j} \frac{d\zeta}{\zeta - w} = 1. \tag{2}$$

If $\gamma_1, \ldots, \gamma_M$ denotes the polygonal curves in (v) of the lemma, then, the cancellations arising from integrating over the same side but in opposite directions in (2) yield

$$\sum_{j=1}^n \frac{1}{2\pi i} \int_{\gamma_j} \frac{d\zeta}{\zeta - w} = 1,$$

and hence $W_{\gamma_{j_0}}(w) \neq 0$ for some j_0. The closed curve γ_{j_0} lies entirely in Ω, and this gives the desired contradiction.

Proof of the lemma. Since F_2 is closed, the sets F_1 and F_2 are at a finite non-zero distance d from one another. Now consider a uniform grid $\mathcal{G}_0$ of the plane consisting of closed squares of side length which is much smaller than d, say $< d/100$, and such that w lies at the center of a closed square R_1 in this grid. Let $\mathcal{R} = \{R_1, \ldots, R_m\}$ denote the finite collection of all closed squares in the grid that intersect F_1. Then, the collection $\mathcal{R}$ satisfies properties (i) through (iv) of the lemma. To guarantee (v), we argue as follows.

The boundary of each square in $\mathcal{R}$ is given the positive (counterclockwise) orientation. The boundary of $\bigcup_{j=1}^{m} R_j$ is then equal to the union of all boundary sides, that is, those sides that do not belong to two adjacent squares in the collection $\mathcal{R}$. Similarly, the boundary vertices are the end-points of all boundary sides. A boundary vertex is said to be "bad," if it is the end-point of more than two boundary sides. (See point P on Figure 1.)

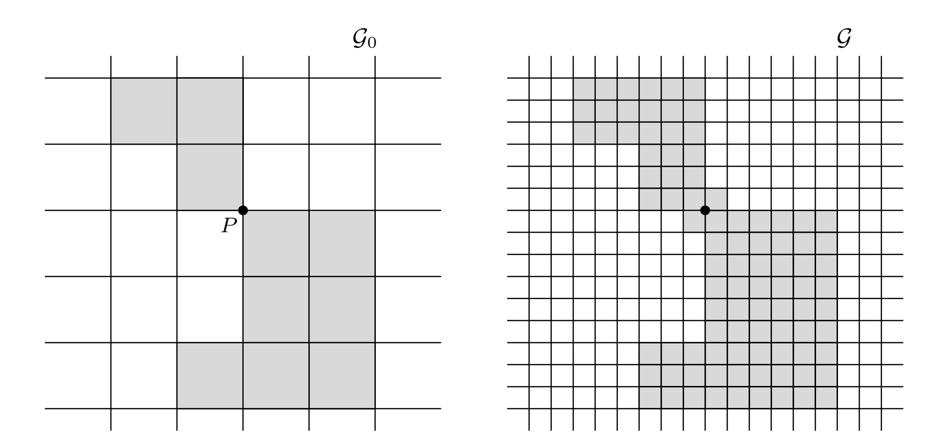


Figure 1. Eliminating bad boundary vertices

To eliminate the bad boundary vertices, we refine the grid $\mathcal{G}_0$ and possibly add some squares. More precisely, consider the grid $\mathcal{G}$ obtained as a refinement of the original grid, by dissecting all squares of $\mathcal{G}_0$ into nine equal subsquares. Then, let $Q_1, \ldots, Q_p$ denote all the squares in the grid $\mathcal{G}$ that are subsquares of squares in the collection $\mathcal{R}$ (so in particular, $p = 9n$), and where Q_1 is chosen so that $w \in Q_1$. Then, we may add finitely many squares from $\mathcal{G}$ near each bad boundary vertex, so that the resulting family $\mathcal{Q} = \{Q_1, \ldots, Q_n\}$ has no bad boundary vertices. (See Figure 1.)

Clearly, $\mathcal{Q}$ still satisfies (i) through (iv), and we claim this collection also satisfies (v). Indeed, let $[a_1, a_2]$ denote any boundary side of $\bigcup_{j=1}^n Q_j$ with its orientation from a_1 to a_2. By considering the three different possibilities, one sees that a_2 is the beginning point of another boundary side $[a_2, a_3]$. Continuing in this fashion, we obtain a sequence of boundary sides $[a_1, a_2], [a_2, a_3], \ldots, [a_n, a_{n+1}], \ldots$. Since there are only finitely many sides, we must have $a_n = a_m$ for some n and some $m > n$. We may choose the smallest m so that $a_n = a_m$, say $m = m'$. Then, we note that if $n > 1$, then $a_{m'}$ is an end-point of at least three boundary sides, namely $[a_{n-1}, a_n]$, $[a_n, a_{n+1}]$, and $[a_{m'-1}, a_{m'}]$, hence $a_{m'}$ is a bad boundary vertex. Since we arranged that $\mathcal{Q}$ had no such boundary vertices, we conclude that $n = 1$, and hence the polygon formed by $a_1, \ldots, a_{m'}$ is closed and simple. We may repeat this process and find that $\mathcal{Q}$ satisfies property (v), and the proof of Lemma 1.5 is complete.

Finally, we are now able to finish the proof of Theorem 1.2, namely, if Ω is bounded and simply connected we can conclude that Ω^c is connected. To see this, note that if Ω^c is not connected, then we have constructed a curve $\gamma \subset \Omega$ and found a point $w \notin \Omega$ so that $W_\gamma(w) \neq 0$, thus contradicting the fact that Ω is simply connected.

2 The Jordan curve theorem

Although we emphasize in the statement of the theorems which follow that the curves are piecewise-smooth, we note that the proofs involve the use of curves that may only be continuous, (the curves Γ_ϵ below).

The two main results in this section are the following.

Theorem 2.1 *Let Γ be curve in the plane that is simple and piecewise-smooth. Then, the complement of Γ is an open connected set whose boundary is precisely Γ.*

Theorem 2.2 *Let Γ be a curve in the plane which is simple, closed, and piecewise-smooth. Then, the complement of Γ consists of two disjoint connected open sets. Precisely one of these regions is bounded and simply connected; it is called the* ***interior*** *of Γ and denoted by Ω. The other component is unbounded, called the* ***exterior*** *of Γ, and denoted by $\mathcal{U}$.*

Moreover, with the appropriate orientation for Γ, we have

$$W_\Gamma(z) = \begin{cases} 1 & \text{if } z \in \Omega, \\ 0 & \text{if } z \in \mathcal{U}. \end{cases}$$

Remark. These two theorems continue to hold in the general case where we drop the assumption that the curves are piecewise-smooth. However, as it turns out, the proofs then are more difficult. Fortunately, the restricted setting of piecewise-smooth curves suffices for many applications.

As a consequence of the above propositions, we may state a version of Cauchy's theorem as follows:

Theorem 2.3 *Suppose f is a function that is holomorphic in the interior Ω of a simple closed curve Γ. Then*

$$\int_\eta f(\zeta)\, d\zeta = 0$$

whenever η is any closed curve contained in Ω.

The idea of the proof of Theorem 2.1 can be roughly summarized as follows. Since the complement of Γ is open, it is sufficient to show it is pathwise connected (Exercise 5, Chapter 1). Let z and w belong to the complement of Γ, and join these two points by a curve. If this curve intersects Γ, we first connect z to z' and w to w', where z' and w' are close to Γ, by curves that do not intersect Γ. Then, we join z' to w' by traveling "parallel" to the curve Γ and going around its end-points if necessary.

Therefore, the key is to construct a family of continuous curves that are "parallel" to Γ. This can be achieved because of the conditions imposed on the curve. Indeed, if γ is a parametrization for a smooth piece of Γ, then γ is continuously differentiable, and $\gamma'(t) \neq 0$. Moreover, the vector $\gamma'(t)$ is tangent to Γ. Consequently, $i\gamma'(t)$ is perpendicular to Γ, and if Γ is simple, considering $\gamma(t) + i\epsilon\gamma'(t)$ amounts to a new curve that is "parallel" to Γ. The details are as follows.

In the next three lemmas and two propositions, we emphasize that Γ_0 denotes a simple *smooth* curve. We recall that an arc-length parametrization γ for a smooth curve Γ_0 satisfies $|\gamma'(t)| = 1$ for all t. Every smooth curve has an arc-length parametrization.

Lemma 2.4 *Let Γ_0 be a simple smooth curve with an arc-length parametrization given by $\gamma : [0, L] \to \mathbb{C}$. For each real number ϵ, let Γ_ϵ be the continuous curve defined by the parametrization*

$$\gamma_\epsilon(t) = \gamma(t) + i\epsilon\gamma'(t), \qquad \textit{for } 0 \leq t \leq L.$$

Then, there exists $\kappa_1 > 0$ so that $\Gamma_0 \cap \Gamma_\epsilon = \emptyset$ whenever $0 < |\epsilon| < \kappa_1$.

Proof. We first prove the result locally. If s and t belong to $[0, L]$, then

$$\begin{aligned}\gamma_\epsilon(t) - \gamma(s) &= \gamma(t) - \gamma(s) + i\epsilon\gamma'(t) \\ &= \int_s^t \gamma'(u)\,du + i\epsilon\gamma'(t) \\ &= \int_s^t [\gamma'(u) - \gamma'(t)]\,du + (t - s + i\epsilon)\gamma'(t).\end{aligned}$$

Since γ' is uniformly continuous on $[0, L]$, there exists $\delta > 0$ so that $|\gamma'(x) - \gamma'(y)| < 1/2$ whenever $|x - y| < \delta$. In particular, if $|s - t| < \delta$ we find that

$$|\gamma_\epsilon(t) - \gamma(s)| > |t - s + i\epsilon|\,|\gamma'(t)| - \frac{|t - s|}{2}.$$

Since γ is an arc-length parametrization, we have $|\gamma'(t)| = 1$, and hence

$$|\gamma_\epsilon(t) - \gamma(s)| > |\epsilon|/2,$$

where we have used the simple fact that $2|a + ib| \geq |a| + |b|$ whenever a and b are real. This proves that $\gamma_\epsilon(t) \neq \gamma(s)$ whenever $|t - s| < \delta$ and $\epsilon \neq 0$.

To conclude the proof of the lemma, we argue as follows. (See Figure 2 for an illustration of the argument.)

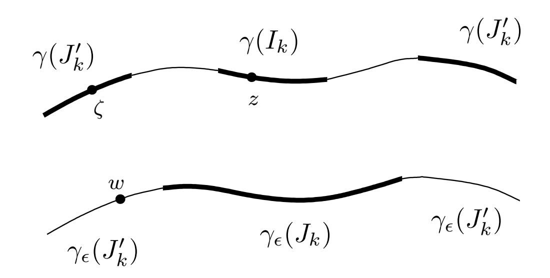


Figure 2. Situation in the proof of Lemma 2.4

Let $0 = t_0 < \cdots < t_n = L$ be a partition of $[0, L]$ with $|t_{k+1} - t_k| < \delta$ for all k, and consider

$$I_k = \{t : |t - t_k| \leq \delta/4\}, \qquad J_k = \{t : |t - t_k| \leq \delta/2\},$$

and

$$J'_k = \{t : |t - t_k| \geq \delta/2\}.$$

Then, we have just proved that

$$\gamma(I_k) \cap \gamma_\epsilon(J_k) = \emptyset \quad \text{whenever } \epsilon \neq 0. \tag{3}$$

Since Γ_0 is simple, the distance d_k between the two compact sets $\gamma(I_k)$ and $\gamma(J'_k)$ is strictly positive. We now claim that

$$\gamma(I_k) \cap \gamma_\epsilon(J'_k) = \emptyset \quad \text{whenever } |\epsilon| < d_k/2. \tag{4}$$

Indeed, if $z \in \gamma(I_k)$ and $w \in \gamma_\epsilon(J'_k)$, then we choose s in J'_k so that $w = \gamma_\epsilon(s)$ and let $\zeta = \gamma(s)$. The triangle inequality then implies

$$|z - w| \geq |z - \zeta| - |\zeta - w| \geq d_k - |\epsilon| \geq d_k/2,$$

and the claim is established. Finally, if we choose $\kappa_1 = \min_k d_k/2$, then (3) and (4) imply that $\Gamma_0 \cap \Gamma_\epsilon = \emptyset$ whenever $0 < |\epsilon| < \kappa_1$, as desired.

The next lemma shows that any point close to an interior point of the curve belongs to one of its parallel translates. By an interior point of the curve, we mean a point of the form $\gamma(t)$ with t in the open interval $(0, L)$. Such a point should not to be confused with an "interior" point of a curve, as in Theorem 2.2.

Lemma 2.5 *Suppose z is a point which does not belong to the smooth curve Γ_0, but that is closer to an interior point of the curve than to either of its end-points. Then z belongs to Γ_ϵ for some $\epsilon \neq 0$.*

More precisely, if $z_0 \in \Gamma_0$ is closest to z and $z_0 = \gamma(t_0)$ for some t_0 in the open interval $(0, L)$, then $z = \gamma(t_0) + i\epsilon\gamma'(t_0)$ for some $\epsilon \neq 0$.

Proof. For t in a neighborhood of t_0 the fact that γ is differentiable guarantees that

$$z - \gamma(t) = z - \gamma(t_0) - \gamma'(t_0)(t - t_0) + o(|t - t_0|).$$

Since $z_0 = \gamma(t_0)$ minimizes the distance from z to Γ_0, we find that

$$\begin{aligned}|z-z_0|^2 \le |z-\gamma(t)|^2 &= |z-z_0|^2 - 2(t-t_0)\mathrm{Re}\left([z-\gamma(t_0)]\overline{\gamma'(t_0)}\right) + \\ &+ o(|t-t_0|).\end{aligned}$$

Since $t-t_0$ can take on positive or negative values, we must have $\mathrm{Re}\left([z-\gamma(t_0)]\overline{\gamma'(t_0)}\right) = 0$, otherwise the above inequality can be violated for t close to t_0. As a result, there exists a real number ϵ with $[z-\gamma(t_0)]\overline{\gamma'(t_0)} = i\epsilon$. Since $|\gamma'(t_0)| = 1$ we have $\overline{\gamma'(t_0)} = 1/\gamma'(t_0)$, and therefore $z-\gamma(t_0) = i\epsilon\gamma'(t_0)$. The proof of the lemma is complete.

Suppose that z and w are close to interior points of Γ_0, so that $z \in \Gamma_\epsilon$ and $w \in \Gamma_\eta$ for some non-zero ϵ and η. If ϵ and η have the same sign, we say that the points z and w belong to the same side of Γ_0. Otherwise, z and w are said to be on opposite sides of Γ_0. We stress the fact that we do not attempt to define the "two sides of Γ_0," but only that given two points near Γ_0, we may infer if they are on the "same side" or on "opposite sides". Also, nothing we have done so far shows that these conditions are mutually exclusive.

Roughly speaking, points on the same side can be joined almost directly by a curve "parallel" to Γ_0, while for points on opposite sides, we also need to go around one of the end-points of Γ_0.

We first investigate the situation for points on the same side of Γ_0.

Proposition 2.6 *Let A and B denote the two end-points of a simple smooth curve Γ_0, and suppose that K is a compact set that satisfies either*

$$\Gamma_0 \cap K = \emptyset \quad \textit{or} \quad \Gamma_0 \cap K = A \cup B.$$

If $z \notin \Gamma_0$ and $w \notin \Gamma_0$ lie on the same side of Γ_0, and are closer to interior points of Γ_0 than they are to K or to the end-points of Γ_0, then z and w can be joined by a continuous curve that lies entirely in the complement of $K \cup \Gamma_0$.

The unspecified compact set K will be chosen appropriately in the proof of the Jordan curve theorem.

Proof. By the previous lemma, consider $z_0 = \gamma(t_0)$ and $w_0 = \gamma(s_0)$ that are interior points of Γ_0 closest to z and w, respectively. Then

$$z = \gamma(t_0) + i\epsilon_0\gamma'(t_0) \quad \text{and} \quad w = \gamma(s_0) + i\eta_0\gamma'(s_0),$$

where ϵ_0 and η_0 have the same sign, which we may assume to be positive. We may also assume that $t_0 \leq s_0$.

The hypothesis of the lemma implies that the line segments joining z to z_0 and w to w_0 are entirely contained in the complement of K and Γ_0. Therefore, for all small $\epsilon > 0$, we may join z and w to the points

$$z_\epsilon = \gamma(t_0) + i\epsilon\gamma'(t_0) \quad \text{and} \quad w = \gamma(s_0) + i\epsilon\gamma'(s_0),$$

respectively. See Figure 3.

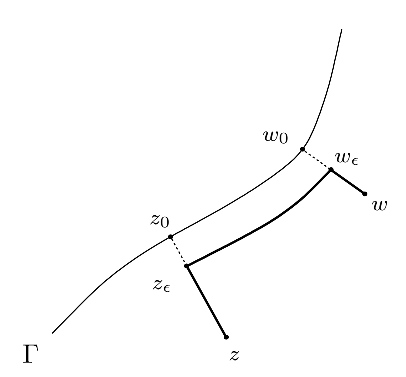


Figure 3. Situation in the proof of Proposition 2.6

Finally, if ϵ is chosen smaller than κ_1 in Lemma 2.4 and also smaller than the distance from K to the part of Γ_0 between z_0 and w_0, that is, $\{\gamma(t) : t_0 \leq t \leq s_0\}$, then the corresponding part of Γ_ϵ, namely $\{\gamma_\epsilon(t) : t_0 \leq t \leq s_0\}$, joins the point z_ϵ to w_ϵ. Moreover, this curve is contained in the complement of K and Γ_0. This proves the proposition.

To join points on opposite sides of Γ_0, we need the following preliminary result, which ensures that there is enough room necessary to travel around the end-points.

Lemma 2.7 *Let Γ_0 be a simple smooth curve. There exists $\kappa_2 > 0$ so that the set N, which consists of points of the form $z = \gamma(L) + \epsilon e^{i\theta}\gamma'(L)$ with $-\pi/2 \leq \theta \leq \pi/2$ and $0 < \epsilon < \kappa_2$, is disjoint from Γ_0.*

Proof. The argument is similar to the one given in the proof of Lemma 2.4. First, we note that

$$\gamma(L) + \epsilon e^{i\theta}\gamma'(L) - \gamma(t) = \int_t^L [\gamma'(u) - \gamma'(L)]\, du + (L - t + \epsilon e^{i\theta})\gamma'(L).$$

If we choose δ so that $|\gamma'(u) - \gamma'(L)| < 1/2$ when $|u - L| < \delta$, then $|t - L| < \delta$ implies

$$|\gamma(L) + \epsilon e^{i\theta}\gamma'(L) - \gamma(t)| \geq |\epsilon|/2\,.$$

Therefore $\gamma(t) \notin N$ whenever $L - \delta \leq t \leq L$. Finally, it suffices to choose κ_2 smaller than the distance from the end-point $\gamma(L)$ to the rest of the curve $\gamma(t)$ with $0 \leq t \leq L - \delta$, to conclude the proof.

Finally, we may state the result analogous to Proposition 2.6 for points that could lie on opposite sides of Γ_0.

Proposition 2.8 *Let A denote an end-point of the simple smooth curve Γ_0, and suppose that K is a compact set that satisfies either*

$$\Gamma_0 \cap K = \emptyset \quad \text{or} \quad \Gamma_0 \cap K = A.$$

If $z \notin \Gamma_0$ and $w \notin \Gamma_0$ are closer to interior points of Γ_0 than they are to K or to the end-points of Γ_0, then z and w can be joined by a continuous curve that lies entirely in the complement of $\Gamma_0 \cup K$.

We only provide an outline of the argument, which is similar to the proof of Proposition 2.6. It suffices to consider the case when z and w lie on opposite sides of Γ_0 and $A = \gamma(0)$. First, we may join

$$z_\epsilon = \gamma(t_0) + i\epsilon\gamma'(t_0) \quad \text{and} \quad w_\epsilon = \gamma(s_0) - i\epsilon\gamma'(s_0)$$

to the points

$$z'_\epsilon = \gamma(L) + i\epsilon\gamma'(L) \quad \text{and} \quad w'_\epsilon = \gamma(L) - i\epsilon\gamma'(L).$$

Then, z'_ϵ and w'_ϵ may be joined within the "half-neighborhood" N of Lemma 2.7. Here, if $t_0 \leq s_0$ we must select $|\epsilon|$ smaller than the distance from $\{\gamma(t) : t_0 \leq t \leq L\}$ to K, and also smaller than κ_1 and κ_2 of Lemmas 2.4 and 2.7.

Proof of Theorem 2.1

Let Γ be a simple piecewise-smooth curve.

First, we prove that the boundary of the set $\mathcal{O} = \Gamma^c$ is precisely Γ. Clearly, $\mathcal{O}$ is an open set whose boundary is contained in Γ. Moreover, any point where Γ is smooth also belongs to the boundary of $\mathcal{O}$ (by Lemma 2.4 for instance). Since the boundary of $\mathcal{O}$ must also be closed, we conclude it is equal to all of Γ.

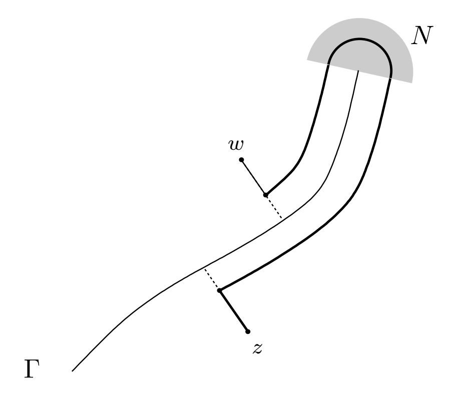


Figure 4. Situation in the proof of Proposition 2.8

The proof that $\mathcal{O}$ is connected is by induction on the number of smooth curves constituting Γ. Suppose first that Γ is simple and smooth, and let Z and W be any two points that do not lie on Γ. Let Λ be any smooth curve in $\mathbb{C}$ that joins Z and W, and which omits the two end-points of Γ. If Λ intersects Γ, it does so at interior points. Therefore, we may join Z by a piece of Λ that does not intersect Γ to a point z that is closer to the interior of Γ than to either of its end-points. Similarly, W can be joined in the complement of Γ to a point w also closer to the interior of Γ than to either of its end-points. Proposition 2.8 (with K empty) then shows that z and w can be joined by a continuous curve in the complement of Γ. Altogether, we may join any two points in the complement of Γ, and this proves the base step of the induction.

Suppose that the theorem is proved for all curves containing $n-1$ smooth curves, and let Γ consist of n smooth curves, so that we may write

$$\Gamma = K \cup \Gamma_0,$$

where K is the union of $n-1$ consecutive smooth curves, and Γ_0 is smooth. In particular, K is compact and intersects Γ_0 in a single one of its end-points. By the induction hypothesis, any two points Z and W in the complement of Γ can be joined by a curve that does not intersect K, and we may also assume that this curve omits both end-points of Γ_0. If this curve intersects Γ_0 in its interior, then we apply Proposition 2.8 to conclude the proof of the theorem.

Proof of Theorem 2.2

Let Γ denote a curve which is simple, closed, and piecewise-smooth. We first prove that the complement of Γ consists of *at most* two components.

Fix a point W that lies outside some large disc that contains Γ, and let $\mathcal{U}$ denote the set of all points that can be joined to W by a continuous curve that lies entirely in the complement of Γ. The set $\mathcal{U}$ is clearly open, and also connected since any two points can be joined by passing first through W. Now we define

$$\Omega = \Gamma^c - \mathcal{U}.$$

We must show that Ω is connected. To this end, let K denote the curve obtained by deleting a smooth piece Γ_0 of Γ. By the Jordan arc theorem, we may join any point $Z \in \Omega$ to W by a curve Λ_Z that does not intersect K. Since $Z \notin \mathcal{U}$, the curve Λ_Z must intersect Γ_0 at one of its interior points. We may therefore choose two points $z, w \in \Lambda_Z$ closer to interior points of Γ_0 than to either of its end-points, and so that the pieces of Λ_Z joining Z to z and W to w are entirely contained in the complement of Γ. Then, the points z and w are on opposite sides of Γ_0, for otherwise, we could apply Proposition 2.6 to find that Z can be joined to W by a curve lying in the complement of Γ, and this contradicts $Z \notin \mathcal{U}$. Finally, if Z_1 is another point in Ω, the two corresponding points z_1 and w_1 must also lie on opposite sides of Γ_0. Moreover, z and z_1 must lie on the same side of Γ_0, for otherwise z and w_1 do, and we can once again join Z to W without crossing Γ, thus contradicting $Z \notin \mathcal{U}$. Therefore, by Proposition 2.6 the points z and z_1 can be joined by a curve in the complement of Γ, and we conclude that Z and Z_1 belong to the same connected component.

The argument thus far proves that Γ^c contains at most two components, but nothing as yet guarantees that Ω is non-empty. To show that Γ^c has precisely two components, it suffices (by Lemma 1.3) to prove that there are points that have different winding numbers with respect to Γ. In fact, we claim that points that are on opposite sides of Γ have winding numbers that differ by 1. To see this, fix a point z_0 on a smooth part of Γ, say $z_0 = \gamma(t_0)$, let $\epsilon > 0$, and define

$$z_\epsilon = \gamma(t_0) + i\epsilon\gamma'(t_0) \quad \text{and} \quad w_\epsilon = \gamma(t_0) - i\epsilon\gamma'(t_0).$$

By our previous observations, points on the same side of Γ belong to the same connected component, and hence

$$\triangle = |W_\Gamma(z_\epsilon) - W_\Gamma(w_\epsilon)|$$

is constant for all small $\epsilon > 0$.

First, we may write

$$\left(\frac{\gamma'(t)}{\gamma(t) - z_\epsilon} - \frac{\gamma'(t)}{\gamma(t) - w_\epsilon}\right) = \frac{2i\epsilon\gamma'(t_0)\gamma'(t)}{[\gamma(t) - \gamma(t_0)]^2 + \epsilon^2\gamma'(t_0)^2}.$$

For the numerator, we use

$$\begin{aligned}\gamma'(t) &= \gamma'(t_0) + [\gamma'(t) - \gamma'(t_0)] \\ &= \gamma'(t_0) + \psi(t),\end{aligned}$$

where $\psi(t) \to 0$ as $t \to t_0$. For the denominator, we recall that $\gamma'(t_0) \neq 0$, so that

$$[\gamma(t) - \gamma(t_0)]^2 + \epsilon^2\gamma'(t_0)^2 = \gamma'(t_0)^2[(t - t_0)^2 + \epsilon^2] + o(|t - t_0|).$$

Putting these results together, we see that

$$\left(\frac{\gamma'(t)}{\gamma(t) - z_\epsilon} - \frac{\gamma'(t)}{\gamma(t) - w_\epsilon}\right) = \frac{2i\epsilon}{(t - t_0)^2 + \epsilon^2} + E(t),$$

where given $\eta > 0$, there exists $\delta > 0$ so that if $|t - t_0| \leq \delta$, the error term satisfies

$$|E(t)| \leq \eta\frac{\epsilon}{(t - t_0)^2 + \epsilon^2}.$$

We then write

$$\triangle = \frac{1}{2\pi i}\int_{|t-t_0|\geq\delta}\left(\frac{\gamma'(t)}{\gamma(t) - z_\epsilon} - \frac{\gamma'(t)}{\gamma(t) - w_\epsilon}\right)dt +$$

$$+\frac{1}{2\pi i}\int_{|t-t_0|<\delta}\left(\frac{2i\epsilon}{(t - t_0)^2 + \epsilon^2} + E(t)\right)dt.$$

The first integral goes to 0 as $\epsilon \to 0$. In the second integral we make the change of variables $t - t_0 = \epsilon s$, and note that

$$\frac{1}{\pi}\int_{-\rho}^{\rho}\frac{ds}{s^2 + 1} = \frac{1}{\pi}\left[\arctan s\right]_{-\rho}^{\rho} \to 1 \qquad \text{as } \rho \to \infty.$$

We therefore see that letting $\epsilon \to 0$ gives

$$|\triangle - 1| < \eta.$$

We conclude that $\triangle = 1$, and hence Γ^c has precisely two components. Finally, only one of these components can be unbounded, namely $\mathcal{U}$, and the winding number of Γ in this component must therefore be zero. By our last result, we see that, after possibly reversing the orientation of the curve, the winding number of any point in the bounded component Ω is constant and equal to 1. Also, it is clear from what has been said that any smooth point on Γ can be approached by points in either component, and hence Γ is the boundary of both Ω and $\mathcal{U}$.

The final step in the proof is to show that the interior of the curve, that is, the bounded component Ω, is simply connected. By Theorem 1.2 it suffices to show that Ω^c is connected. If not, then

$$\Omega^c = F_1 \cup F_2,$$

where F_1 and F_2 are closed, disjoint, and non-empty. Let

$$\mathcal{O}_1 = \mathcal{U} \cap F_1 \quad \text{and} \quad \mathcal{O}_2 = \mathcal{U} \cap F_2.$$

Clearly, $\mathcal{O}_1$ and $\mathcal{O}_2$ are disjoint. If $z \in \mathcal{O}_1$, then $z \in \mathcal{U}$, and every small ball centered at z is contained in $\mathcal{U}$. If every such ball intersects F_2, then $z \in F_2$ since F_2 is closed. However, F_1 and F_2 are disjoint, so this cannot happen. Consequently, $\mathcal{O}_1$ is open, and by the same argument, so is $\mathcal{O}_2$. Finally, we claim that $\mathcal{O}_1$ is non-empty. If not, then F_1 is entirely contained in Γ and $\mathcal{U}$ is contained in F_2. Pick any point $z \in F_1$, which we know belongs to Γ. Now every ball centered at z intersects $\mathcal{U}$, hence F_2. But F_2 is closed and disjoint from F_1, so we get a contradiction. A similar argument for $\mathcal{O}_1$ proves that

$$\mathcal{U} = \mathcal{O}_1 \cup \mathcal{O}_2,$$

where $\mathcal{O}_1, \mathcal{O}_2$ are disjoint, open, and non-empty. This contradicts the fact that $\mathcal{U}$ is connected, and concludes the proof of the Jordan curve theorem for piecewise-smooth curves.

2.1 Proof of a general form of Cauchy's theorem

Theorem 2.9 *If a function f is holomorphic in an open set that contains a simple closed piecewise-smooth curve Γ and its interior, then*

$$\int_\Gamma f = 0.$$

Let $\mathcal{O}$ denote an open set on which f is holomorphic, and which contains Γ and its interior Ω. The idea is to construct a closed curve Λ

in Ω that is so close to Γ that $\int_\Gamma f = \int_\Lambda f$. Then, the integral on the right-hand side is 0, since f is holomorphic in the simply connected open set Ω. We build Λ as follows. Near the smooth parts of Γ, the curve Λ is essentially a curve like Γ_ϵ in Lemma 2.4. Near points where smooth parts of Γ join, we shall use for Λ an arc of a circle. This is illustrated in Figure 5.

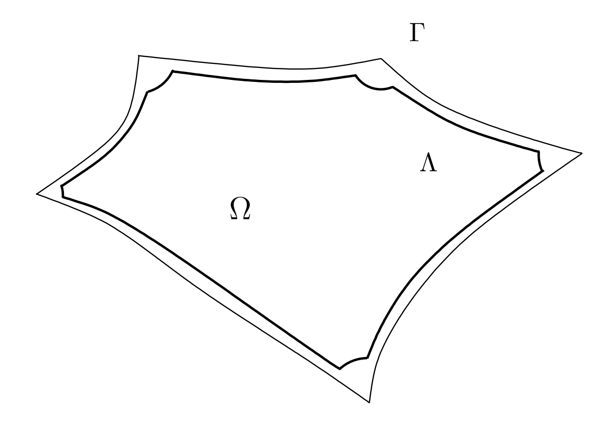


Figure 5. The curve Λ

To find the appropriate connecting arcs, we need the following preliminary result.

Lemma 2.10 *Let $\gamma : [0,1] \to \mathbb{C}$ be a simple smooth curve. Then, for all sufficiently small $\delta > 0$ the circle C_δ centered at $\gamma(0)$ and of radius δ intersects γ in precisely one point.*

Proof. We may assume that $\gamma(0) = 0$. Since $\gamma(0) \neq \gamma(1)$ it is clear that for each small $\delta > 0$, the circle C_δ intersects γ in at least one point. If the conclusion in the lemma is false, we can find a sequence of positive δ_j going to 0, and so that the equation $|\gamma(t)| = \delta_j$ has at least two distinct solutions. The mean value theorem applied to $h(t) = |\gamma(t)|^2$ provides a sequence of positive numbers t_j so that $t_j \to 0$ and $h'(t_j) = 0$. Thus

$$\gamma'(t_j) \cdot \gamma(t_j) = 0 \qquad \text{for all } j.$$

However, the curve is smooth, so

$$\gamma(t) = \gamma(0) + \gamma'(0)t + t\varphi(t) \qquad \text{and} \qquad \gamma'(t) = \gamma'(0) + \psi(t),$$

where $|\varphi(t)| \to 0$ and $|\psi(t)| \to 0$ as t goes to 0. Then recalling that $\gamma(0) = 0$, we find $\gamma'(t) \cdot \gamma(t) = |\gamma'(0)|^2 t + o(|t|)$. The definition of a smooth curve also requires that $\gamma'(0) \neq 0$, so the above gives

$$\gamma'(t) \cdot \gamma(t) \neq 0 \qquad \text{for all small } t.$$

This is the desired contradiction.

Returning to the proof of Cauchy's theorem, choose ϵ so small that the open set $\mathcal{U}$ of all points at a distance $< \epsilon$ of Γ is contained in $\mathcal{O}$.

Next, if $P_1, \ldots, P_n$ denote the consecutive points where smooth parts of Γ join, we may pick $\delta < \epsilon/10$ so small that each circle C_j centered at a point P_j and of radius δ intersects Γ in precisely two distinct points (this is possible by the previous lemma). These two points on C_j determine two arcs of circles, only one of which (denoted by $\mathcal{C}_j$) has an interior entirely contained in Ω. To see this, it suffices to recall that if γ is a parametrization of a smooth part of Γ with end-point P_j, then for all small ϵ' the curves parametrized by $\gamma_{\epsilon'}$ and $\gamma_{-\epsilon'}$ of Lemma 2.4 lie on opposite sides of Γ and must intersect the circle C_j. By construction the disc D_j^* centered at P_j and of radius 2δ is also contained in $\mathcal{U}$, hence in $\mathcal{O}$.

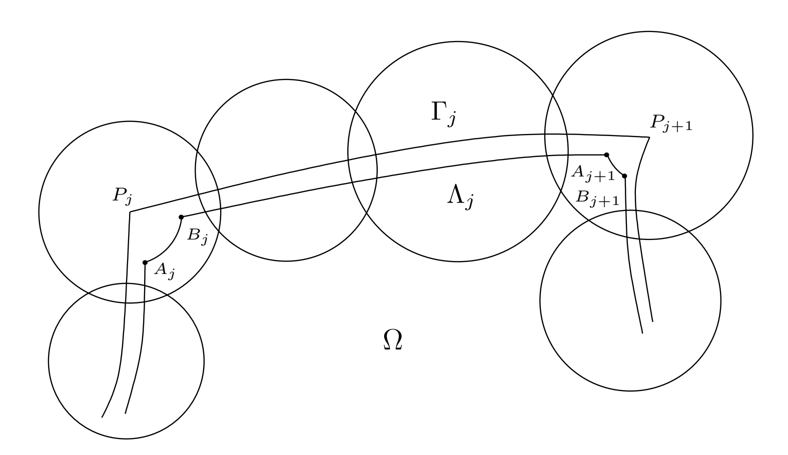


Figure 6. Construction of the curve Λ

We wish to construct Λ so that we may argue as in the proof of Theorem 5.1, Chapter 3 and establish $\int_\Gamma f = \int_\Lambda f$. To do so, we consider a chain of discs $\mathcal{D} = \{D_0, \ldots, D_K\}$ contained in $\mathcal{U}$, and so that Γ is contained in their union, with $D_k \cap D_{k+1} \neq \emptyset$, $D_0 = D_K$, and with the discs

D_j^* part of the chain $\mathcal{D}$. Suppose Γ_j is the smooth part of Γ that joins P_j to P_{j+1}. By Lemma 2.4 it is possible to construct a continuous curve Λ_j that is contained in Ω and in the union of the discs, and which connects a point on B_j on $\mathcal{C}_j$ to a point on A_{j+1} on $\mathcal{C}_{j+1}$ (see Figure 6). Since we only assumed that Γ has one continuous derivative, Λ_j need not be smooth, but by approximating this continuous curve by polygonal lines if necessary, we may actually assume that Λ_j is also smooth. Then, A_{j+1} is joined to B_{j+1} by a piece of $\mathcal{C}_{j+1}$, and so on. This procedure provides a piecewise-smooth curve Λ that is closed and contained in Ω.

Since f has a primitive on each disc of the family $\mathcal{D}$, we may argue as in the proof of Theorem 5.1, Chapter 3 to find that $\int_\Gamma f = \int_\Lambda f$. Since Ω is simply connected, we have $\int_\Lambda f = 0$, and as a result

$$\int_\Gamma f = 0.$$

Notes and References

Useful references for many of the subjects treated here are Saks and Zygmund [34], Ahlfors [2], and Lang [23].

Introduction

The citation is from Riemann's dissertation [32].

Chapter 1

The citation is a free translation of a passage in Borel's book [6].

Chapter 2

The citation is a translation of an excerpt from Cauchy's memoir [7].

Results related to the natural boundaries of holomorphic functions in the unit disc can be found in Titchmarsh [36].

The construction of the universal functions in Problem 5 are due to G. D. Birkhoff and G.R. MacLane.

Chapter 3

The citation is a translation of a passage in Cauchy's memoir [8].

Problem 1 and other results related to injective holomorphic mappings (univalent functions) can be found in Duren [11].

Also, see Muskhelishvili [25] for more about the Cauchy integral introduced in Problem 5.

Chapter 4

The citation is from Wiener [40].

The argument in Exercise 1 was discovered by D. J. Newman; see [4].

The Paley-Wiener theorems appeared first in [28]; further generalizations can be found in Stein and Weiss [35].

Results related to the Borel transform (Problem 4) can be found in Boas [5].

Chapter 5

The citation is a translation from the German of a passage in a letter from K. Weierstrass to S. Kowalewskaja; see [38].

A classical reference for Nevanlinna theory is the book by R. Nevanlinna himself [27].

Chapter 6

A number of different proofs of the analytic continuation and functional equation for the zeta function can be found in Chapter 2 of Titchmarsch [37].

Chapter 7

The citation is from Hadamard [14]. Riemann's statement concerning the zeroes of the zeta function in the critical strip is a passage taken from his paper [33].

Further material related to the proof of the prime number theorem presented in the text is in Chapter 2 of Ingham [19], and Chapter 3 of Titchmarsch [37].

The "elementary" analysis of the distribution of primes (without using the analytic properties of the zeta function) was initiated by Tchebychev, and culminated in the Erdös-Selberg proof of the prime number theorem. See Chapter XXII in Hardy and Wright [17].

The results in Problems 2 and 3 can be found in Chapter 4 of Ingham [19].

For Problem 4, consult Estermann [13].

Chapter 8

The citation is from Christoffel [9].

A systematic treatment of conformal mappings is Nehari [26].

Some history related to the Riemann mapping theorem, as well as the details in Problem 7, can be found in Remmert [31].

Results related to the boundary behavior of holomorphic functions (Problem 6) are in Chapter XIV of Zygmund [41].

An introduction to the interplay between the Poincaré metric and complex analysis can be found in Ahlfors [1]. For further results on the Schwartz-Pick lemma and hyperbolicity, see Kobayashi [21].

For more on Bieberbach's conjecture, see Chapter 2 in Duren [11] and Chapter 8 in Hayman [18].

Chapter 9

The citation is taken from Poincaré [30].

Problems 2, 3, and 4 are in Saks and Zygmund [34].

Chapter 10

The citation is from Hardy, Chapter IX in [16].

A systematic account of the theory of theta functions and Jacobi's theory of elliptic functions is in Whittaker and Watson [39], Chapters 21 and 22.

Section 2. For more on the partition function, see Chapter XIX in Hardy and Wright [17].

Section 3. The more standard proofs of the theorems about the sum of two and four squares are in Hardy and Wright [17], Chapter XX. The approach we use was developed by Mordell and Hardy [15] to derive exact formulas for the number of representations as the sum of k squares, when $k \geq 5$. The special case $k = 8$ is in Problem 6. For $k \leq 4$ the method as given there breaks down because of the non-absolute convergence of the associated "Eisenstein series." In our presentation we get around this difficulty by using the "forbidden" Eisenstein series. When $k = 2$, an entirely different construction is needed, and the analysis centering around $\mathcal{C}(\tau)$ is a further new aspect of this problem.

The theorem on the sum of three squares (Problem 1) is in Part I, Chapter 4 of Landau [22].

Appendix A

The citation is taken from the appendix in Airy's article [3].

For systematic accounts of Laplace's method, stationary phase, and the method of steepest descent, see Erdélyi [12] and Copson [10].

The more refined asymptotics of the partition function can be found in Chapter 8 of Hardy [16].

Appendix B

The citation is taken from Picard's address found in Jordan's collected works [20].

The proof of the Jordan curve theorem for piecewise-smooth curve due to Pederson [29] is an adaptation of the proof for polygonal curves which can be found in Saks and Zygmund [34].

For a proof of the Jordan theorem for continuous curves using notions of algebraic topology, see Munkres [24].

Bibliography

[1] L. V. Ahlfors. *Conformal Invariants*. McGraw-Hill, New York, 1973.

[2] L. V. Ahlfors. *Complex Analysis*. McGraw-Hill, New York, third edition, 1979.

[3] G. B. Airy. On the intensity of light in the neighbourhood of a caustic. *Transactions of the Cambridge Philosophical Society*, 6:379–402, 1838.

[4] J. Bak and D. J. Newman. *Complex Analysis*. Springer-Verlag, New York, second edition, 1997.

[5] R. P. Boas. *Entire Functions*. Academic Press, New York, 1954.

[6] E. Borel. *L'imaginaire et le réel en Mathématiques et en Physique*. Albin Michel, Paris, 1952.

[7] A. L. Cauchy. Mémoires sur les intégrales définies. *Oeuvres complètes d'Augustin Cauchy, Gauthier-Villars, Paris*, Iere Série(I), 1882.

[8] A. L. Cauchy. Sur un nouveau genre de calcul analogue au calcul infinitesimal. *Oeuvres complètes d'Augustin Cauchy, Gauthier-Villars, Paris*, IIeme Série(VI), 1887.

[9] E. B. Christoffel. Ueber die Abbildung einer Einblättrigen, Einfach Zusammenhägenden, Ebenen Fläche auf Einem Kreise. *Nachrichten von der Königl. Gesselschaft der Wissenchaft und der G. A. Universität zu Göttingen*, pages 283–298, 1870.

[10] E.T. Copson. *Asymptotic Expansions*, volume 55 of *Cambridge Tracts in Math. and Math Physics*. Cambridge University Press, 1965.

[11] P. L. Duren. *Univalent Functions*. Springer-Verlag, New York, 1983.

[12] A. Erdélyi. *Asymptotic Expansions*. Dover, New York, 1956.

[13] T. Estermann. *Introduction to Modern Prime Number Theory*. Cambridge University Press, 1952.

[14] J. Hadamard. *The Psychology of Invention in the Mathematical Field*. Princeton University Press, 1945.

[15] G. H. Hardy. On the representation of a number as the sum of any number of squares, and in particular five. *Trans. Amer. Math. Soc*, 21:255–284, 1920.

[16] G. H. Hardy. *Ramanujan.* Cambridge University Press, 1940.

[17] G. H. Hardy and E. M. Wright. *An introduction to the Theory of Numbers.* Oxford University Press, London, fifth edition, 1979.

[18] W. K. Hayman. *Multivalent Functions.* Cambridge University Press, second edition, 1994.

[19] A. E. Ingham. *The Distribution of Prime Numbers.* Cambridge University Press, 1990.

[20] C. Jordan. *Oeuvres de Camille Jordan*, volume IV. Gauthier-Villars, Paris, 1964.

[21] S. Kobayashi. *Hyperbolic Manifolds and Holomorphic Mappings.* M. Dekker, New York, 1970.

[22] E. Landau. *Vorlesungen über Zahlentheorie*, volume 1. S. Hirzel, Leipzig, 1927.

[23] S. Lang. *Complex Analysis.* Springer-Verlag, New York, fourth edition, 1999.

[24] J. R. Munkres. *Elements of Algebraic Topology.* Addison-Wesley, Reading, MA, 1984.

[25] N. I. Muskhelishvili. *Singular Integral Equations.* Noordhott International Publishing, Leyden, 1977.

[26] Z. Nehari. *Comformal Mapping.* McGraw-Hill, New York, 1952.

[27] R. Nevanlinna. *Analytic functions.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellung. Springer-Verlag, New York, 1970.

[28] R. Paley and N. Wiener. *Fourier Transforms in the Complex Domain*, volume XIX of *Colloquium publications.* American Mathematical Society, Providence, RI, 1934.

[29] R. N. Pederson. The Jordan curve theorem for piecewise smooth curves. *Amer. Math. Monthly*, 76:605–610, 1969.

[30] H. Poincaré. L'Oeuvre mathématiques de Weierstrass. *Acta Mathematica*, 22, 1899.

[31] R. Remmert. *Classical Topics in Complex Function Theory.* Springer-Verlag, New York, 1998.

[32] B. Riemann. Grundlagen für eine Allgemeine Theorie der Functionen einer Veränderlichen Complexen Grösse. *Inauguraldissertation, Göttingen, 1851*, Collected Works, Springer-Verlag, 1990.

[33] B. Riemann. Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monat. Preuss. Akad. Wissen., 1859*, Collected Works, Springer-Verlag, 1990.

[34] S. Saks and Z. Zygmund. *Analytic Functions.* Elsevier, PWN-Polish Scientific, third edition, 1971.

[35] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, 1971.

[36] E. C. Titchmarsh. *The Theory of Functions.* Oxford University Press, London, second edition, 1939.

[37] E. C. Titchmarsh. *The Theory of the Riemann Zeta-Function.* Oxford University Press, 1951.

[38] K. Weierstrass. *Briefe von Karl Weierstrass an Sofie Kowalewskaja 1871-1891.* Moskva, Nauka, 1973.

[39] E.T. Whittaker and G.N. Watson. *A Course in Modern Analysis.* Cambridge University Press, 1927.

[40] N. Wiener. "R. E. A. C. Paley - in Memoriam". *Bull. Amer. Math. Soc.*, 39:476, 1933.

[41] A. Zygmund. *Trigonometric Series*, volume I and II. Cambridge University Press, second edition, 1959. Reprinted 1993.

Symbol Glossary

The page numbers on the right indicate the first time the symbol or notation is defined or used. As usual, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ denote the integers, the rationals, the reals, and the complex numbers respectively.

Symbol	Meaning	Page
$\mathrm{Re}(z)$, $\mathrm{Im}(z)$	Real and Imaginary parts	2
$\arg z$	Argument of z	4
$\lvert z\rvert$, $\overline{z}$	Absolute value and complex conjugate	3, 3
$D_r(z_0)$, $\overline{D}_r(z_0)$	Open and closed discs centered at z_0 and with radius r	5, 6
$C_r(z_0)$	Circle centered at z_0 with radius r	6
D, C	Generic disc and circle	
$\mathbb{D}$	Unit disc	6
Ω^c, $\overline{\Omega}$, $\partial\Omega$	Complement, closure, and boundary of Ω	6
$\mathrm{diam}(\Omega)$	Diameter of Ω	6
$\frac{\partial}{\partial z}$, $\frac{\partial}{\partial \overline{z}}$	Differential operators	12
e^z, $\cos z$, $\sin z$	Complex exponential and trigonometric functions	14, 16
γ^-	Reverse parametrization	19
O, o, $\sim$	Bounds and asymptotic relations	24
$\triangle$	Laplacian	27
$F(\alpha,\beta,\gamma;z)$	Hypergeometric series	28
$\mathrm{res}_z f$	Residue	75
$P_r(\gamma)$, $\mathcal{P}_y(x)$	Poisson kernels	67, 78
$\cosh z$, $\sinh z$	Hyperbolic cosine and sine	81, 83
$\mathbb{S}$	Riemann sphere	89
$\log$, $\log_\Omega$	Logarithms	98, 99
$\hat{f}(\xi)$	Fourier transform	111
$\mathfrak{F}_a$, $\mathfrak{F}$	Class of functions with moderate decay in strips	113, 114
S_a, $S_{\delta,M}$	Horizontal strips	113, 160
ρ, ρ_f	Order of growth	138
E_k	Canonical factors	145
ψ_α	Blaschke factors	153

$\Gamma(s)$	Gamma function	160
$\zeta(s)$	Riemann zeta function	168
ϑ, $\Theta(z\|\tau)$, $\theta(\tau)$	Theta function	169, 284, 284
$\xi(s)$	Xi function	169
J_ν	Bessel functions	176
B_m	Bernoulli number	179
$\pi(x)$	Number of primes $\leq x$	182
$f(x) \approx g(x)$	Asymptotic relation	182
$\psi(x)$, $\Lambda(n)$, $\psi_1(x)$	Functions of Tchebychev	188, 189, 190
$d(n)$	Number of divisors of n	200
$\sigma_a(n)$	Sum of the ath powers of divisors of n	200
$\mu(n)$	Möbius function	200
$\text{Li}(x)$	Approximation to $\pi(x)$	202
$\mathbb{H}$	Upper half-plane	208
$\text{Aut}(\Omega)$	Automorphism group of Ω	219
$\text{SL}_2(\mathbb{R})$	Special linear group	222
$\text{PSL}_2(\mathbb{R})$	Projective special linear group	223
$\text{SU}(1,1)$	Group of fractional linear transformations	257
Λ, Λ^*	Lattice and lattice minus the origin	262, 267
$\wp$	Weierstrass elliptic function	269
$E_k(\tau)$, $E_2^*(\tau)$	Eisenstein series	273, 305
$F(\tau)$, $\tilde{F}(\tau)$	Forbidden Eisenstein series and its reverse	278, 305
$\Pi(z\|\tau)$	Triple product	286
$\eta(\tau)$	Dedekind eta function	292
$p(n)$	Partition function	293
$r_2(n)$	Number of ways n is a sum of two squares	296
$r_4(n)$	Number of ways n is a sum of four squares	297
$d_1(n)$, $d_3(n)$, $\sigma_1^*(n)$	Divisor functions	297, 304
$\text{Ai}(s)$	Airy function	328
$W_\gamma(z)$	Winding number	347

Index

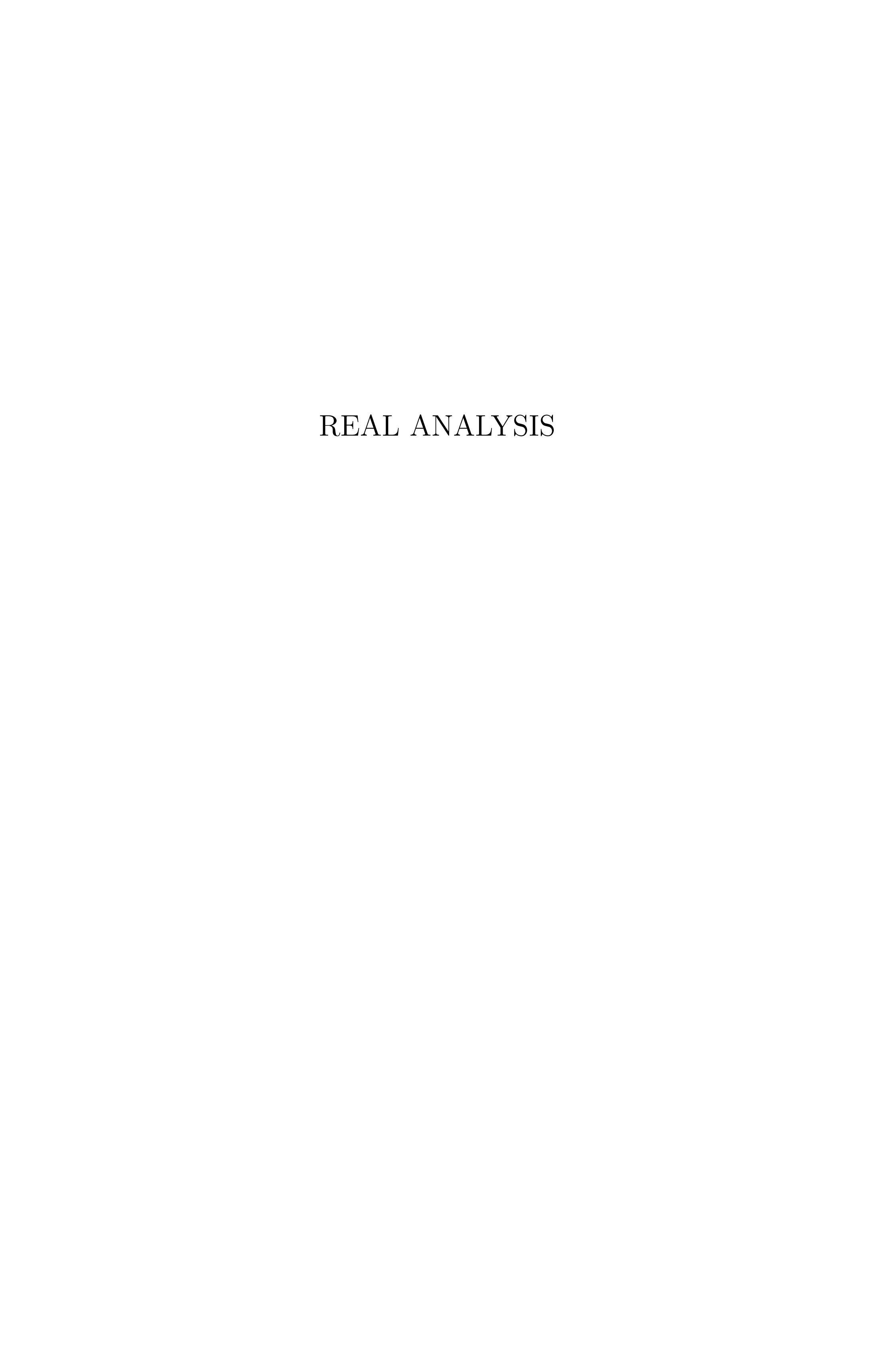
Relevant items that also arise in Book I are listed in this index, preceeded by the numeral I.

PRINCETON LECTURES IN ANALYSIS III

REAL ANALYSIS

MEASURE THEORY, INTEGRATION, & HILBERT SPACES

ELIAS M. STEIN & RAMI SHAKARCHI



REAL ANALYSIS

Princeton Lectures in Analysis

I Fourier Analysis: An Introduction

II Complex Analysis

III Real Analysis:
Measure Theory, Integration, and
Hilbert Spaces

IV Functional Analysis: Introduction
to Further Topics in Analysis

Princeton Lectures in Analysis

III

REAL ANALYSIS

Measure Theory, Integration, and Hilbert Spaces

Elias M. Stein

&

Rami Shakarchi

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

Library of Congress Control Number 2004114065
ISBN 978-0-691-11386-9

British Library Cataloging-in-Publication Data is available

The publisher would like to acknowledge the authors of this volume for
providing the camera-ready copy from which this book was printed

Printed on acid-free paper. ∞

press.princeton.edu

Printed in the United States of America

3 5 7 9 10 8 6 4

To my grandchildren

Carolyn, Alison, Jason

E.M.S.

To my parents

Mohamed & Mireille

and my brother

Karim

R.S.

Foreword

Beginning in the spring of 2000, a series of four one-semester courses were taught at Princeton University whose purpose was to present, in an integrated manner, the core areas of analysis. The objective was to make plain the organic unity that exists between the various parts of the subject, and to illustrate the wide applicability of ideas of analysis to other fields of mathematics and science. The present series of books is an elaboration of the lectures that were given.

While there are a number of excellent texts dealing with individual parts of what we cover, our exposition aims at a different goal: presenting the various sub-areas of analysis not as separate disciplines, but rather as highly interconnected. It is our view that seeing these relations and their resulting synergies will motivate the reader to attain a better understanding of the subject as a whole. With this outcome in mind, we have concentrated on the main ideas and theorems that have shaped the field (sometimes sacrificing a more systematic approach), and we have been sensitive to the historical order in which the logic of the subject developed.

We have organized our exposition into four volumes, each reflecting the material covered in a semester. Their contents may be broadly summarized as follows:

I. Fourier series and integrals.

II. Complex analysis.

III. Measure theory, Lebesgue integration, and Hilbert spaces.

IV. A selection of further topics, including functional analysis, distributions, and elements of probability theory.

However, this listing does not by itself give a complete picture of the many interconnections that are presented, nor of the applications to other branches that are highlighted. To give a few examples: the elements of (finite) Fourier series studied in Book I, which lead to Dirichlet characters, and from there to the infinitude of primes in an arithmetic progression; the X-ray and Radon transforms, which arise in a number of

problems in Book I, and reappear in Book III to play an important role in understanding Besicovitch-like sets in two and three dimensions; Fatou's theorem, which guarantees the existence of boundary values of bounded holomorphic functions in the disc, and whose proof relies on ideas developed in each of the first three books; and the theta function, which first occurs in Book I in the solution of the heat equation, and is then used in Book II to find the number of ways an integer can be represented as the sum of two or four squares, and in the analytic continuation of the zeta function.

A few further words about the books and the courses on which they were based. These courses where given at a rather intensive pace, with 48 lecture-hours a semester. The weekly problem sets played an indispensable part, and as a result exercises and problems have a similarly important role in our books. Each chapter has a series of "Exercises" that are tied directly to the text, and while some are easy, others may require more effort. However, the substantial number of hints that are given should enable the reader to attack most exercises. There are also more involved and challenging "Problems"; the ones that are most difficult, or go beyond the scope of the text, are marked with an asterisk.

Despite the substantial connections that exist between the different volumes, enough overlapping material has been provided so that each of the first three books requires only minimal prerequisites: acquaintance with elementary topics in analysis such as limits, series, differentiable functions, and Riemann integration, together with some exposure to linear algebra. This makes these books accessible to students interested in such diverse disciplines as mathematics, physics, engineering, and finance, at both the undergraduate and graduate level.

It is with great pleasure that we express our appreciation to all who have aided in this enterprise. We are particularly grateful to the students who participated in the four courses. Their continuing interest, enthusiasm, and dedication provided the encouragement that made this project possible. We also wish to thank Adrian Banner and José Luis Rodrigo for their special help in running the courses, and their efforts to see that the students got the most from each class. In addition, Adrian Banner also made valuable suggestions that are incorporated in the text.

We wish also to record a note of special thanks for the following individuals: Charles Fefferman, who taught the first week (successfully launching the whole project!); Paul Hagelstein, who in addition to reading part of the manuscript taught several weeks of one of the courses, and has since taken over the teaching of the second round of the series; and Daniel Levine, who gave valuable help in proof-reading. Last but not least, our thanks go to Gerree Pecht, for her consummate skill in typesetting and for the time and energy she spent in the preparation of all aspects of the lectures, such as transparencies, notes, and the manuscript.

We are also happy to acknowledge our indebtedness for the support we received from the 250th Anniversary Fund of Princeton University, and the National Science Foundation's VIGRE program.

Elias M. Stein
Rami Shakarchi

Princeton, New Jersey
August 2002

In this third volume we establish the basic facts concerning measure theory and integration. This allows us to reexamine and develop further several important topics that arose in the previous volumes, as well as to introduce a number of other subjects of substantial interest in analysis. To aid the interested reader, we have starred sections that contain more advanced material. These can be omitted on first reading. We also want to take this opportunity to thank Daniel Levine for his continuing help in proof-reading and the many suggestions he made that are incorporated in the text.

November 2004

Contents

Introduction

> I turn away in fright and horror from this lamentable plague of functions that do not have derivatives.
>
> *C. Hermite,* 1893

Starting in about 1870 a revolutionary change in the conceptual framework of analysis began to take shape, one that ultimately led to a vast transformation and generalization of the understanding of such basic objects as functions, and such notions as continuity, differentiability, and integrability.

The earlier view that the relevant functions in analysis were given by formulas or other "analytic" expressions, that these functions were by their nature continuous (or nearly so), that by necessity such functions had derivatives for most points, and moreover these were integrable by the accepted methods of integration − all of these ideas began to give way under the weight of various examples and problems that arose in the subject, which could not be ignored and required new concepts to be understood. Parallel with these developments came new insights that were at once both more geometric and more abstract: a clearer understanding of the nature of curves, their rectifiability and their extent; also the beginnings of the theory of sets, starting with subsets of the line, the plane, etc., and the "measure" that could be assigned to each.

That is not to say that there was not considerable resistance to the change of point-of-view that these advances required. Paradoxically, some of the leading mathematicians of the time, those who should have been best able to appreciate the new departures, were among the ones who were most skeptical. That the new ideas ultimately won out can be understood in terms of the many questions that could now be addressed. We shall describe here, somewhat imprecisely, several of the most significant such problems.

1 Fourier series: completion

Whenever f is a (Riemann) integrable function on $[-\pi, \pi]$ we defined in Book I its Fourier series $f \sim \sum a_n e^{inx}$ by

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx, \tag{1}$$

and saw then that one had Parseval's identity,

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx.$$

However, the above relationship between functions and their Fourier coefficients is not completely reciprocal when limited to Riemann integrable functions. Thus if we consider the space $\mathcal{R}$ of such functions with its square norm, and the space $\ell^2(\mathbb{Z})$ with its norm,[1] each element f in $\mathcal{R}$ assigns a corresponding element $\{a_n\}$ in $\ell^2(\mathbb{Z})$, and the two norms are identical. However, it is easy to construct elements in $\ell^2(\mathbb{Z})$ that do not correspond to functions in $\mathcal{R}$. Note also that the space $\ell^2(\mathbb{Z})$ is *complete* in its norm, while $\mathcal{R}$ is not.[2] Thus we are led to two questions:

(i) What are the putative "functions" f that arise when we complete $\mathcal{R}$? In other words: given an arbitrary sequence $\{a_n\} \in \ell^2(\mathbb{Z})$ what is the nature of the (presumed) function f corresponding to these coefficients?

(ii) How do we integrate such functions f (and in particular verify (1))?

2 Limits of continuous functions

Suppose $\{f_n\}$ is a sequence of continuous functions on $[0, 1]$. We assume that $\lim_{n\to\infty} f_n(x) = f(x)$ exists for every x, and inquire as to the nature of the limiting function f.

If we suppose that the convergence is uniform, matters are straightforward and f is then everywhere continuous. However, once we drop the assumption of uniform convergence, things may change radically and the issues that arise can be quite subtle. An example of this is given by the fact that one can construct a sequence of continuous functions $\{f_n\}$ converging everywhere to f so that

[1] We use the notation of Chapter 3 in Book I.

[2] See the discussion surrounding Theorem 1.1 in Section 1, Chapter 3 of Book I.

(a) $0 \leq f_n(x) \leq 1$ for all x.

(b) The sequence $f_n(x)$ is montonically decreasing as $n \to \infty$.

(c) The limiting function f is not Riemann integrable.[3]

However, in view of (a) and (b), the sequence $\int_0^1 f_n(x)\,dx$ converges to a limit. So it is natural to ask: what method of integration can be used to integrate f and obtain that for it

$$\int_0^1 f(x)\,dx = \lim_{n\to\infty} \int_0^1 f_n(x)\,dx\,?$$

It is with Lebesgue integration that we can solve both this problem and the previous one.

3 Length of curves

The study of curves in the plane and the calculation of their lengths are among the first issues dealt with when one learns calculus. Suppose we consider a continuous curve Γ in the plane, given parametrically by $\Gamma = \{(x(t), y(t))\}$, $a \leq t \leq b$, with x and y continuous functions of t. We define the *length* of Γ in the usual way: as the supremum of the lengths of all polygonal lines joining successively finitely many points of Γ, taken in order of increasing t. We say that Γ is *rectifiable* if its length L is finite. When $x(t)$ and $y(t)$ are continuously differentiable we have the well-known formula,

$$L = \int_a^b \left((x'(t))^2 + (y'(t))^2\right)^{1/2}\,dt. \tag{2}$$

The problems we are led to arise when we consider general curves. More specifically, we can ask:

(i) What are the conditions on the functions $x(t)$ and $y(t)$ that guarantee the rectifiability of Γ?

(ii) When these are satisfied, does the formula (2) hold?

The first question has a complete answer in terms of the notion of functions of "bounded variation." As to the second, it turns out that if x and y are of bounded variation, the integral (2) is always meaningful; however, the equality fails in general, but can be restored under appropriate reparametrization of the curve Γ.

[3]The limit f can be highly discontinuous. See, for instance, Exercise 10 in Chapter 1.

There are further issues that arise. Rectifiable curves, because they are endowed with length, are genuinely one-dimensional in nature. Are there (non-rectifiable) curves that are two-dimensional? We shall see that, indeed, there are continuous curves in the plane that fill a square, or more generally have any dimension between 1 and 2, if the notion of fractional dimension is appropriately defined.

4 Differentiation and integration

The so-called "fundamental theorem of the calculus" expresses the fact that differentiation and integration are inverse operations, and this can be stated in two different ways, which we abbreviate as follows:

$$F(b) - F(a) = \int_a^b F'(x)\, dx, \tag{3}$$

$$\frac{d}{dx} \int_0^x f(y)\, dy = f(x). \tag{4}$$

For the first assertion, the existence of continuous functions F that are nowhere differentiable, or for which $F'(x)$ exists for every x, but F' is not integrable, leads to the problem of finding a general class of the F for which (3) is valid. As for (4), the question is to formulate properly and establish this assertion for the general class of integrable functions f that arise in the solution of the first two problems considered above. These questions can be answered with the help of certain "covering" arguments, and the notion of absolute continuity.

5 The problem of measure

To put matters clearly, the fundamental issue that must be understood in order to try to answer all the questions raised above is the problem of measure. Stated (imprecisely) in its version in two dimensions, it is the problem of assigning to each subset E of $\mathbb{R}^2$ its two-dimensional measure $m_2(E)$, that is, its "area," extending the standard notion defined for elementary sets. Let us instead state more precisely the analogous problem in one dimension, that of constructing one-dimensional measure $m_1 = m$, which generalizes the notion of length in $\mathbb{R}$.

We are looking for a non-negative function m defined on the family of subsets E of $\mathbb{R}$ that we allow to be extended-valued, that is, to take on the value $+\infty$. We require:

(a) $m(E) = b - a$ if E is the interval $[a, b]$, $a \le b$, of length $b - a$.

(b) $m(E) = \sum_{n=1}^{\infty} m(E_n)$ whenever $E = \bigcup_{n=1}^{\infty} E_n$ and the sets E_n are disjoint.

Condition (b) is the "countable additivity" of the measure m. It implies the special case:

(b′) $m(E_1 \cup E_2) = m(E_1) + m(E_2)$ if E_1 and E_2 are disjoint.

However, to apply the many limiting arguments that arise in the theory the general case (b) is indispensable, and (b′) by itself would definitely be inadequate.

To the axioms (a) and (b) one adds the translation-invariance of m, namely

(c) $m(E + h) = m(E)$, for every $h \in \mathbb{R}$.

A basic result of the theory is the existence (and uniqueness) of such a measure, Lebesgue measure, when one limits oneself to a class of reasonable sets, those which are "measurable." This class of sets is closed under countable unions, intersections, and complements, and contains the open sets, the closed sets, and so forth.[4]

It is with the construction of this measure that we begin our study. From it will flow the general theory of integration, and in particular the solutions of the problems discussed above.

A chronology

We conclude this introduction by listing some of the signal events that marked the early development of the subject.

1872 – Weierstrass's construction of a nowhere differentiable function.

1881 – Introduction of functions of bounded variation by Jordan and later (1887) connection with rectifiability.

1883 – Cantor's ternary set.

1890 – Construction of a space-filling curve by Peano.

1898 – Borel's measurable sets.

1902 – Lebesgue's theory of measure and integration.

1905 – Construction of non-measurable sets by Vitali.

1906 – Fatou's application of Lebesgue theory to complex analysis.

[4]There is no such measure on the class of all subsets, since there exist non-measurable sets. See the construction of such a set at the end of Section 3, Chapter 1.

1 Measure Theory

> The sets whose measure we can define by virtue of the preceding ideas we will call measurable sets; we do this without intending to imply that it is not possible to assign a measure to other sets.
>
> *E. Borel,* 1898

This chapter is devoted to the construction of Lebesgue measure in $\mathbb{R}^d$ and the study of the resulting class of measurable functions. After some preliminaries we pass to the first important definition, that of exterior measure for any subset E of $\mathbb{R}^d$. This is given in terms of approximations by unions of cubes that cover E. With this notion in hand we can define measurability and thus restrict consideration to those sets that are measurable. We then turn to the fundamental result: the collection of measurable sets is closed under complements and countable unions, and the measure is additive if the subsets in the union are disjoint.

The concept of measurable functions is a natural outgrowth of the idea of measurable sets. It stands in the same relation as the concept of continuous functions does to open (or closed) sets. But it has the important advantage that the class of measurable functions is closed under pointwise limits.

1 Preliminaries

We begin by discussing some elementary concepts which are basic to the theory developed below.

The main idea in calculating the "volume" or "measure" of a subset of $\mathbb{R}^d$ consists of approximating this set by unions of other sets whose geometry is simple and whose volumes are known. It is convenient to speak of "volume" when referring to sets in $\mathbb{R}^d$; but in reality it means "area" in the case $d = 2$ and "length" in the case $d = 1$. In the approach given here we shall use rectangles and cubes as the main building blocks of the theory: in $\mathbb{R}$ we use intervals, while in $\mathbb{R}^d$ we take products of intervals. In all dimensions rectangles are easy to manipulate and have a standard notion of volume that is given by taking the product of the length of all sides.

Next, we prove two simple theorems that highlight the importance of these rectangles in the geometry of open sets: in $\mathbb{R}$ every open set is a countable union of disjoint open intervals, while in $\mathbb{R}^d$, $d \geq 2$, every open set is "almost" the disjoint union of closed cubes, in the sense that only the boundaries of the cubes can overlap. These two theorems motivate the definition of exterior measure given later.

We shall use the following standard notation. A **point** $x \in \mathbb{R}^d$ consists of a d-tuple of real numbers

$$x = (x_1, x_2, \ldots, x_d), \qquad x_i \in \mathbb{R}, \text{ for } i = 1, \ldots, d.$$

Addition of points is componentwise, and so is multiplication by a real scalar. The **norm** of x is denoted by $|x|$ and is defined to be the standard Euclidean norm given by

$$|x| = \left(x_1^2 + \cdots + x_d^2\right)^{1/2}.$$

The **distance** between two points x and y is then simply $|x - y|$.

The **complement** of a set E in $\mathbb{R}^d$ is denoted by E^c and defined by

$$E^c = \{x \in \mathbb{R}^d : x \notin E\}.$$

If E and F are two subsets of $\mathbb{R}^d$, we denote the complement of F in E by

$$E - F = \{x \in \mathbb{R}^d : x \in E \text{ and } x \notin F\}.$$

The **distance** between two sets E and F is defined by

$$d(E, F) = \inf |x - y|,$$

where the infimum is taken over all $x \in E$ and $y \in F$.

Open, closed, and compact sets

The **open ball** in $\mathbb{R}^d$ centered at x and of radius r is defined by

$$B_r(x) = \{y \in \mathbb{R}^d : |y - x| < r\}.$$

A subset $E \subset \mathbb{R}^d$ is **open** if for every $x \in E$ there exists $r > 0$ with $B_r(x) \subset E$. By definition, a set is **closed** if its complement is open.

We note that any (not necessarily countable) union of open sets is open, while in general the intersection of only finitely many open sets

is open. A similar statement holds for the class of closed sets, if one interchanges the roles of unions and intersections.

A set E is **bounded** if it is contained in some ball of finite radius. A bounded set is **compact** if it is also closed. Compact sets enjoy the Heine-Borel covering property:

- Assume E is compact, $E \subset \bigcup_\alpha \mathcal{O}_\alpha$, and each $\mathcal{O}_\alpha$ is open. Then there are finitely many of the open sets, $\mathcal{O}_{\alpha_1}, \mathcal{O}_{\alpha_2}, \ldots, \mathcal{O}_{\alpha_N}$, such that $E \subset \bigcup_{j=1}^N \mathcal{O}_{\alpha_j}$.

In words, *any* covering of a compact set by a collection of open sets contains a *finite* subcovering.

A point $x \in \mathbb{R}^d$ is a **limit point** of the set E if for every $r > 0$, the ball $B_r(x)$ contains points of E. This means that there are points in E which are arbitrarily close to x. An **isolated point** of E is a point $x \in E$ such that there exists an $r > 0$ where $B_r(x) \cap E$ is equal to $\{x\}$.

A point $x \in E$ is an **interior point** of E if there exists $r > 0$ such that $B_r(x) \subset E$. The set of all interior points of E is called the **interior** of E. Also, the **closure** $\overline{E}$ of the E consists of the union of E and all its limit points. The **boundary** of a set E, denoted by ∂E, is the set of points which are in the closure of E but not in the interior of E.

Note that the closure of a set is a closed set; every point in E is a limit point of E; and a set is closed if and only if it contains all its limit points. Finally, a closed set E is **perfect** if E does not have any isolated points.

Rectangles and cubes

A (closed) **rectangle** R in $\mathbb{R}^d$ is given by the product of d one-dimensional closed and bounded intervals

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d],$$

where $a_j \leq b_j$ are real numbers, $j = 1, 2, \ldots, d$. In other words, we have

$$R = \{(x_1, \ldots, x_d) \in \mathbb{R}^d : \ a_j \leq x_j \leq b_j \quad \text{for all } j = 1, 2, \ldots, d\}.$$

We remark that in our definition, a rectangle is *closed* and has sides *parallel* to the coordinate axis. In $\mathbb{R}$, the rectangles are precisely the closed and bounded intervals, while in $\mathbb{R}^2$ they are the usual four-sided rectangles. In $\mathbb{R}^3$ they are the closed parallelepipeds.

We say that the lengths of the sides of the rectangle R are $b_1 - a_1, \ldots, b_d - a_d$. The **volume** of the rectangle R is denoted by $|R|$, and

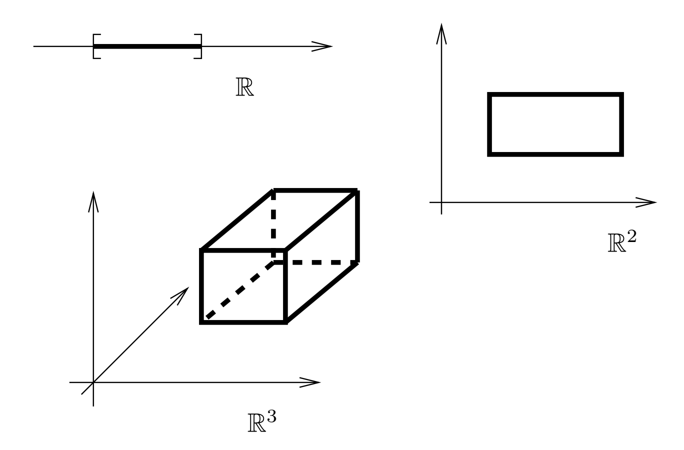


Figure 1. Rectangles in $\mathbb{R}^d$, $d = 1, 2, 3$

is defined to be

$$|R| = (b_1 - a_1) \cdots (b_d - a_d).$$

Of course, when $d = 1$ the "volume" equals length, and when $d = 2$ it equals area.

An open rectangle is the product of open intervals, and the interior of the rectangle R is then

$$(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d).$$

Also, a **cube** is a rectangle for which $b_1 - a_1 = b_2 - a_2 = \cdots = b_d - a_d$. So if $Q \subset \mathbb{R}^d$ is a cube of common side length ℓ, then $|Q| = \ell^d$.

A union of rectangles is said to be **almost disjoint** if the interiors of the rectangles are disjoint.

In this chapter, coverings by rectangles and cubes play a major role, so we isolate here two important lemmas.

Lemma 1.1 *If a rectangle is the almost disjoint union of finitely many other rectangles, say* $R = \bigcup_{k=1}^{N} R_k$, *then*

$$|R| = \sum_{k=1}^{N} |R_k|.$$

Proof. We consider the grid formed by extending indefinitely the sides of all rectangles $R_1, \ldots, R_N$. This construction yields finitely many rectangles $\tilde{R}_1, \ldots, \tilde{R}_M$, and a partition $J_1, \ldots, J_N$ of the integers between 1 and M, such that the unions

$$R = \bigcup_{j=1}^{M} \tilde{R}_j \quad \text{and} \quad R_k = \bigcup_{j \in J_k} \tilde{R}_j, \quad \text{for } k = 1, \ldots, N$$

are almost disjoint (see the illustration in Figure 2).

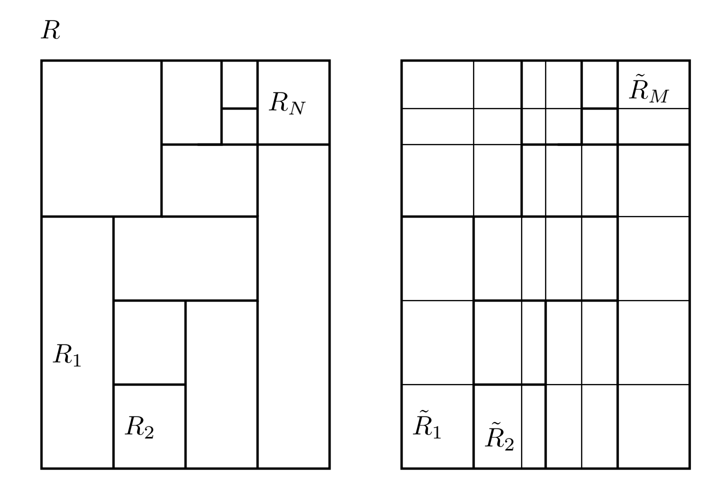


Figure 2. The grid formed by the rectangles R_k

For the rectangle R, for example, we see that $|R| = \sum_{j=1}^{M} |\tilde{R}_j|$, since the grid actually partitions the sides of R and each $\tilde{R}_j$ consists of taking products of the intervals in these partitions. Thus when adding the volumes of the $\tilde{R}_j$ we are summing the corresponding products of lengths of the intervals that arise. Since this also holds for the other rectangles $R_1, \ldots, R_N$, we conclude that

$$|R| = \sum_{j=1}^{M} |\tilde{R}_j| = \sum_{k=1}^{N} \sum_{j \in J_k} |\tilde{R}_j| = \sum_{k=1}^{N} |R_k|.$$

A slight modification of this argument then yields the following:

Lemma 1.2 *If $R, R_1, \ldots, R_N$ are rectangles, and $R \subset \bigcup_{k=1}^N R_k$, then*

$$|R| \leq \sum_{k=1}^N |R_k|.$$

The main idea consists of taking the grid formed by extending all sides of the rectangles $R, R_1, \ldots, R_N$, and noting that the sets corresponding to the J_k (in the above proof) need not be disjoint any more.

We now proceed to give a description of the structure of open sets in terms of cubes. We begin with the case of $\mathbb{R}$.

Theorem 1.3 *Every open subset $\mathcal{O}$ of $\mathbb{R}$ can be writen uniquely as a countable union of disjoint open intervals.*

Proof. For each $x \in \mathcal{O}$, let I_x denote the largest open interval containing x and contained in $\mathcal{O}$. More precisely, since $\mathcal{O}$ is open, x is contained in some small (non-trivial) interval, and therefore if

$$a_x = \inf\{a < x : (a, x) \subset \mathcal{O}\} \quad \text{and} \quad b_x = \sup\{b > x : (x, b) \subset \mathcal{O}\}$$

we must have $a_x < x < b_x$ (with possibly infinite values for a_x and b_x). If we now let $I_x = (a_x, b_x)$, then by construction we have $x \in I_x$ as well as $I_x \subset \mathcal{O}$. Hence

$$\mathcal{O} = \bigcup_{x \in \mathcal{O}} I_x.$$

Now suppose that two intervals I_x and I_y intersect. Then their union (which is also an open interval) is contained in $\mathcal{O}$ and contains x. Since I_x is maximal, we must have $(I_x \cup I_y) \subset I_x$, and similarly $(I_x \cup I_y) \subset I_y$. This can happen only if $I_x = I_y$; therefore, any two distinct intervals in the collection $\mathcal{I} = \{I_x\}_{x \in \mathcal{O}}$ must be disjoint. The proof will be complete once we have shown that there are only countably many distinct intervals in the collection $\mathcal{I}$. This, however, is easy to see, since every open interval I_x contains a rational number. Since different intervals are disjoint, they must contain distinct rationals, and therefore $\mathcal{I}$ is countable, as desired.

Naturally, if $\mathcal{O}$ is open and $\mathcal{O} = \bigcup_{j=1}^\infty I_j$, where the I_j's are disjoint open intervals, the measure of $\mathcal{O}$ ought to be $\sum_{j=1}^\infty |I_j|$. Since this representation is unique, we could take this as a definition of measure; we would then note that whenever $\mathcal{O}_1$ and $\mathcal{O}_2$ are open and disjoint, the measure of their union is the sum of their measures. Although this provides

a natural notion of measure for an open set, it is not immediately clear how to generalize it to other sets in $\mathbb{R}$. Moreover, a similar approach in higher dimensions already encounters complications even when defining measures of open sets, since in this context the direct analogue of Theorem 1.3 is not valid (see Exercise 12). There is, however, a substitute result.

Theorem 1.4 *Every open subset $\mathcal{O}$ of $\mathbb{R}^d$, $d \geq 1$, can be written as a countable union of almost disjoint closed cubes.*

Proof. We must construct a countable collection $\mathcal{Q}$ of closed cubes whose interiors are disjoint, and so that $\mathcal{O} = \bigcup_{Q \in \mathcal{Q}} Q$.

As a first step, consider the grid in $\mathbb{R}^d$ formed by taking all closed cubes of side length 1 whose vertices have integer coordinates. In other words, we consider the natural grid of lines parallel to the axes, that is, the grid generated by the lattice $\mathbb{Z}^d$. We shall also use the grids formed by cubes of side length 2^{-N} obtained by successively bisecting the original grid.

We either accept or reject cubes in the initial grid as part of $\mathcal{Q}$ according to the following rule: if Q is entirely contained in $\mathcal{O}$ then we accept Q; if Q intersects both $\mathcal{O}$ and $\mathcal{O}^c$ then we tentatively accept it; and if Q is entirely contained in $\mathcal{O}^c$ then we reject it.

As a second step, we bisect the tentatively accepted cubes into 2^d cubes with side length $1/2$. We then repeat our procedure, by accepting the smaller cubes if they are completely contained in $\mathcal{O}$, tentatively accepting them if they intersect both $\mathcal{O}$ and $\mathcal{O}^c$, and rejecting them if they are contained in $\mathcal{O}^c$. Figure 3 illustrates these steps for an open set in $\mathbb{R}^2$.

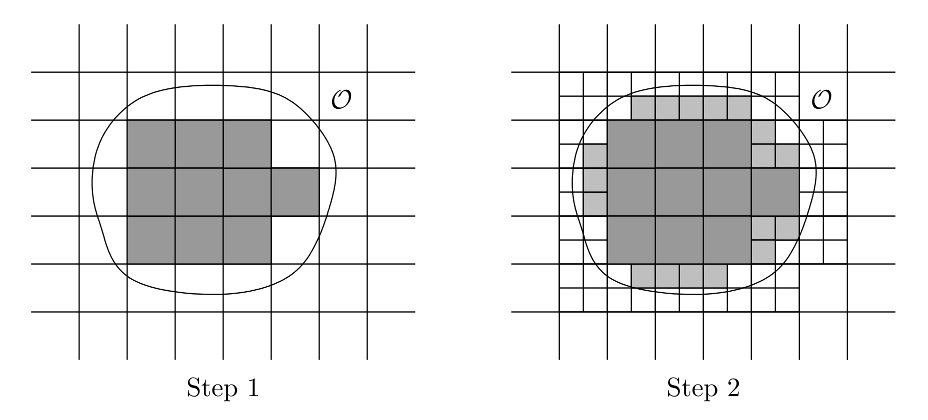


Figure 3. Decomposition of $\mathcal{O}$ into almost disjoint cubes

This procedure is then repeated indefinitely, and (by construction) the resulting collection $\mathcal{Q}$ of all accepted cubes is countable and consists of almost disjoint cubes. To see why their union is all of $\mathcal{O}$, we note that given $x \in \mathcal{O}$ there exists a cube of side length 2^{-N} (obtained from successive bisections of the original grid) that contains x and that is entirely contained in $\mathcal{O}$. Either this cube has been accepted, or it is contained in a cube that has been previously accepted. This shows that the union of all cubes in $\mathcal{Q}$ covers $\mathcal{O}$.

Once again, if $\mathcal{O} = \bigcup_{j=1}^{\infty} R_j$ where the rectangles R_j are almost disjoint, it is reasonable to assign to $\mathcal{O}$ the measure $\sum_{j=1}^{\infty} |R_j|$. This is natural since the volume of the boundary of each rectangle should be 0, and the overlap of the rectangles should not contribute to the volume of $\mathcal{O}$. We note, however, that the above decomposition into cubes is not unique, and it is not immediate that the sum is independent of this decomposition. So in $\mathbb{R}^d$, with $d \geq 2$, the notion of volume or area, even for open sets, is more subtle.

The general theory developed in the next section actually yields a notion of volume that is consistent with the decompositions of open sets of the previous two theorems, and applies to all dimensions. Before we come to that, we discuss an important example in $\mathbb{R}$.

The Cantor set

The Cantor set plays a prominent role in set theory and in analysis in general. It and its variants provide a rich source of enlightening examples.

We begin with the closed unit interval $C_0 = [0, 1]$ and let C_1 denote the set obtained from deleting the middle third open interval from $[0, 1]$, that is,

$$C_1 = [0, 1/3] \cup [2/3, 1].$$

Next, we repeat this procedure for each sub-interval of C_1; that is, we delete the middle third open interval. At the second stage we get

$$C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1].$$

We repeat this process for each sub-interval of C_2, and so on (Figure 4).

This procedure yields a sequence C_k, $k = 0, 1, 2, \ldots$ of compact sets with

$$C_0 \supset C_1 \supset C_2 \supset \cdots \supset C_k \supset C_{k+1} \supset \cdots .$$

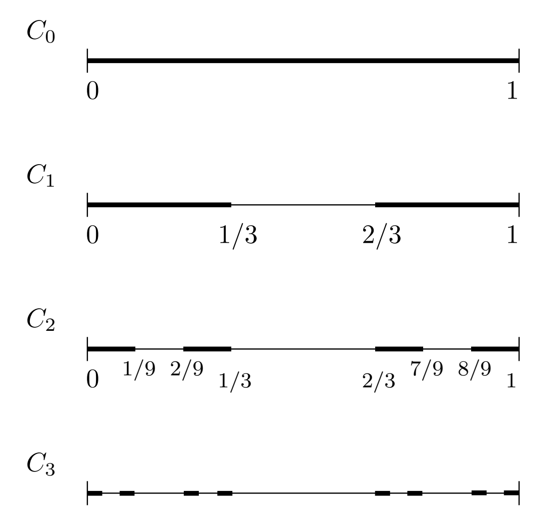


Figure 4. Construction of the Cantor set

The **Cantor set** $\mathcal{C}$ is by definition the intersection of all C_k's:

$$\mathcal{C} = \bigcap_{k=0}^{\infty} C_k.$$

The set $\mathcal{C}$ is not empty, since all end-points of the intervals in C_k (all k) belong to $\mathcal{C}$.

Despite its simple construction, the Cantor set enjoys many interesting topological and analytical properties. For instance, $\mathcal{C}$ is closed and bounded, hence compact. Also, $\mathcal{C}$ is totally disconnected: given any $x, y \in \mathcal{C}$ there exists $z \notin \mathcal{C}$ that lies between x and y. Finally, $\mathcal{C}$ is perfect: it has no isolated points (Exercise 1).

Next, we turn our attention to the question of determining the "size" of $\mathcal{C}$. This is a delicate problem, one that may be approached from different angles depending on the notion of size we adopt. For instance, in terms of cardinality the Cantor set is rather large: it is not countable. Since it can be mapped to the interval $[0, 1]$, the Cantor set has the cardinality of the continuum (Exercise 2).

However, from the point of view of "length" the size of $\mathcal{C}$ is small. Roughly speaking, the Cantor set has length zero, and this follows from the following intuitive argument: the set $\mathcal{C}$ is covered by sets C_k whose lengths go to zero. Indeed, C_k is a disjoint union of 2^k intervals of length

3^{-k}, making the total length of C_k equal to $(2/3)^k$. But $\mathcal{C} \subset C_k$ for all k, and $(2/3)^k \to 0$ as k tends to infinity. We shall define a notion of measure and make this argument precise in the next section.

2 The exterior measure

The notion of exterior measure is the first of two important concepts needed to develop a theory of measure. We begin with the definition and basic properties of exterior measure. Loosely speaking, the exterior measure m_* assigns to *any* subset of $\mathbb{R}^d$ a first notion of size; various examples show that this notion coincides with our earlier intuition. However, the exterior measure lacks the desirable property of additivity when taking the union of disjoint sets. We remedy this problem in the next section, where we discuss in detail the other key concept of measure theory, the notion of measurable sets.

The exterior measure, as the name indicates, attempts to describe the volume of a set E by approximating it from the outside. The set E is covered by cubes, and if the covering gets finer, with fewer cubes overlapping, the volume of E should be close to the sum of the volumes of the cubes.

The precise definition is as follows: if E is *any* subset of $\mathbb{R}^d$, the **exterior measure**[1] of E is

$$m_*(E) = \inf \sum_{j=1}^{\infty} |Q_j|, \tag{1}$$

where the infimum is taken over all countable coverings $E \subset \bigcup_{j=1}^{\infty} Q_j$ by closed cubes. The exterior measure is always non-negative but could be infinite, so that in general we have $0 \leq m_*(E) \leq \infty$, and therefore takes values in the extended positive numbers.

We make some preliminary remarks about the definition of the exterior measure given by (1).

(i) It is important to note that it would not suffice to allow *finite* sums in the definition of $m_*(E)$. The quantity that would be obtained if one considered only coverings of E by finite unions of cubes is in general larger than $m_*(E)$. (See Exercise 14.)

(ii) One can, however, replace the coverings by cubes, with coverings by rectangles; or with coverings by balls. That the former alternative

[1] Some authors use the term **outer measure** instead of exterior measure.

yields the same exterior measure is quite direct. (See Exercise 15.) The equivalence with the latter is more subtle. (See Exercise 26 in Chapter 3.)

We begin our investigation of this new notion by providing examples of sets whose exterior measures can be calculated, and we check that the latter matches our intuitive idea of volume (length in one dimension, area in two dimensions, etc.)

EXAMPLE 1. The exterior measure of a point is zero. This is clear once we observe that a point is a cube with volume zero, and which covers itself. Of course the exterior measure of the empty set is also zero.

EXAMPLE 2. The exterior measure of a closed cube is equal to its volume. Indeed, suppose Q is a closed cube in $\mathbb{R}^d$. Since Q covers itself, we must have $m_*(Q) \leq |Q|$. Therefore, it suffices to prove the reverse inequality.

We consider an arbitrary covering $Q \subset \bigcup_{j=1}^{\infty} Q_j$ by cubes, and note that it suffices to prove that

$$|Q| \leq \sum_{j=1}^{\infty} |Q_j|. \tag{2}$$

For a fixed $\epsilon > 0$ we choose for each j an open cube S_j which contains Q_j, and such that $|S_j| \leq (1+\epsilon)|Q_j|$. From the open covering $\bigcup_{j=1}^{\infty} S_j$ of the compact set Q, we may select a finite subcovering which, after possibly renumbering the rectangles, we may write as $Q \subset \bigcup_{j=1}^{N} S_j$. Taking the closure of the cubes S_j, we may apply Lemma 1.2 to conclude that $|Q| \leq \sum_{j=1}^{N} |S_j|$. Consequently,

$$|Q| \leq (1+\epsilon)\sum_{j=1}^{N} |Q_j| \leq (1+\epsilon)\sum_{j=1}^{\infty} |Q_j|.$$

Since ϵ is arbitrary, we find that the inequality (2) holds; thus $|Q| \leq m_*(Q)$, as desired.

EXAMPLE 3. If Q is an open cube, the result $m_*(Q) = |Q|$ still holds. Since Q is covered by its closure $\overline{Q}$, and $|\overline{Q}| = |Q|$, we immediately see that $m_*(Q) \leq |Q|$. To prove the reverse inequality, we note that if Q_0 is a closed cube contained in Q, then $m_*(Q_0) \leq m_*(Q)$, since any covering of Q by a countable number of closed cubes is also a covering of Q_0 (see Observation 1 below). Hence $|Q_0| \leq m_*(Q)$, and since we can choose Q_0 with a volume as close as we wish to $|Q|$, we must have $|Q| \leq m_*(Q)$.

EXAMPLE 4. The exterior measure of a rectangle R is equal to its volume. Indeed, arguing as in Example 2, we see that $|R| \leq m_*(R)$. To obtain the reverse inequality, consider a grid in $\mathbb{R}^d$ formed by cubes of side length $1/k$. Then, if $\mathcal{Q}$ consists of the (finite) collection of all cubes entirely contained in R, and $\mathcal{Q}'$ the (finite) collection of all cubes that intersect the complement of R, we first note that $R \subset \bigcup_{Q \in (\mathcal{Q} \cup \mathcal{Q}')} Q$. Also, a simple argument yields

$$\sum_{Q \in \mathcal{Q}} |Q| \leq |R|.$$

Moreover, there are $O(k^{d-1})$ cubes[2] in $\mathcal{Q}'$, and these cubes have volume k^{-d}, so that $\sum_{Q \in \mathcal{Q}'} |Q| = O(1/k)$. Hence

$$\sum_{Q \in (\mathcal{Q} \cup \mathcal{Q}')} |Q| \leq |R| + O(1/k),$$

and letting k tend to infinity yields $m_*(R) \leq |R|$, as desired.

EXAMPLE 5. The exterior measure of $\mathbb{R}^d$ is infinite. This follows from the fact that any covering of $\mathbb{R}^d$ is also a covering of any cube $Q \subset \mathbb{R}^d$, hence $|Q| \leq m_*(\mathbb{R}^d)$. Since Q can have arbitrarily large volume, we must have $m_*(\mathbb{R}^d) = \infty$.

EXAMPLE 6. The Cantor set $\mathcal{C}$ has exterior measure 0. From the construction of $\mathcal{C}$, we know that $\mathcal{C} \subset C_k$, where each C_k is a disjoint union of 2^k closed intervals, each of length 3^{-k}. Consequently, $m_*(\mathcal{C}) \leq (2/3)^k$ for all k, hence $m_*(\mathcal{C}) = 0$.

Properties of the exterior measure

The previous examples and comments provide some intuition underlying the definition of exterior measure. Here, we turn to the further study of m_* and prove five properties of exterior measure that are needed in what follows.

First, we record the following remark that is immediate from the definition of m_*:

---

[2]We remind the reader of the notation $f(x) = O(g(x))$, which means that $|f(x)| \leq C|g(x)|$ for some constant C and all x in a given range. In this particular example, there are fewer than Ck^{d-1} cubes in question, as $k \to \infty$.

- For every $\epsilon > 0$, there exists a covering $E \subset \bigcup_{j=1}^{\infty} Q_j$ with

$$\sum_{j=1}^{\infty} m_*(Q_j) \leq m_*(E) + \epsilon.$$

The relevant properties of exterior measure are listed in a series of observations.

Observation 1 (Monotonicity) *If $E_1 \subset E_2$, then $m_*(E_1) \leq m_*(E_2)$.*

This follows once we observe that any covering of E_2 by a countable collection of cubes is also a covering of E_1.

In particular, monotonicity implies that every bounded subset of $\mathbb{R}^d$ has finite exterior measure.

Observation 2 (Countable sub-additivity) *If $E = \bigcup_{j=1}^{\infty} E_j$, then $m_*(E) \leq \sum_{j=1}^{\infty} m_*(E_j)$.*

First, we may assume that each $m_*(E_j) < \infty$, for otherwise the inequality clearly holds. For any $\epsilon > 0$, the definition of the exterior measure yields for each j a covering $E_j \subset \bigcup_{k=1}^{\infty} Q_{k,j}$ by closed cubes with

$$\sum_{k=1}^{\infty} |Q_{k,j}| \leq m_*(E_j) + \frac{\epsilon}{2^j}.$$

Then, $E \subset \bigcup_{j,k=1}^{\infty} Q_{k,j}$ is a covering of E by closed cubes, and therefore

$$\begin{aligned} m_*(E) \leq \sum_{j,k} |Q_{k,j}| &= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |Q_{k,j}| \\ &\leq \sum_{j=1}^{\infty} \left(m_*(E_j) + \frac{\epsilon}{2^j} \right) \\ &= \sum_{j=1}^{\infty} m_*(E_j) + \epsilon. \end{aligned}$$

Since this holds true for every $\epsilon > 0$, the second observation is proved.

Observation 3 *If $E \subset \mathbb{R}^d$, then $m_*(E) = \inf m_*(\mathcal{O})$, where the infimum is taken over all open sets $\mathcal{O}$ containing E.*

By monotonicity, it is clear that the inequality $m_*(E) \leq \inf m_*(\mathcal{O})$ holds. For the reverse inequality, let $\epsilon > 0$ and choose cubes Q_j such that $E \subset \bigcup_{j=1}^{\infty} Q_j$, with

$$\sum_{j=1}^{\infty} |Q_j| \leq m_*(E) + \frac{\epsilon}{2}.$$

Let Q_j^0 denote an open cube containing Q_j, and such that $|Q_j^0| \leq |Q_j| + \epsilon/2^{j+1}$. Then $\mathcal{O} = \bigcup_{j=1}^{\infty} Q_j^0$ is open, and by Observation 2

$$\begin{aligned} m_*(\mathcal{O}) \leq \sum_{j=1}^{\infty} m_*(Q_j^0) &= \sum_{j=1}^{\infty} |Q_j^0| \\ &\leq \sum_{j=1}^{\infty} \left(|Q_j| + \frac{\epsilon}{2^{j+1}} \right) \\ &\leq \sum_{j=1}^{\infty} |Q_j| + \frac{\epsilon}{2} \\ &\leq m_*(E) + \epsilon. \end{aligned}$$

Hence $\inf m_*(\mathcal{O}) \leq m_*(E)$, as was to be shown.

Observation 4 *If $E = E_1 \cup E_2$, and $d(E_1, E_2) > 0$, then*

$$m_*(E) = m_*(E_1) + m_*(E_2).$$

By Observation 2, we already know that $m_*(E) \leq m_*(E_1) + m_*(E_2)$, so it suffices to prove the reverse inequality. To this end, we first select δ such that $d(E_1, E_2) > \delta > 0$. Next, we choose a covering $E \subset \bigcup_{j=1}^{\infty} Q_j$ by closed cubes, with $\sum_{j=1}^{\infty} |Q_j| \leq m_*(E) + \epsilon$. We may, after subdividing the cubes Q_j, assume that each Q_j has a diameter less than δ. In this case, each Q_j can intersect at most one of the two sets E_1 or E_2. If we denote by J_1 and J_2 the sets of those indices j for which Q_j intersects E_1 and E_2, respectively, then $J_1 \cap J_2$ is empty, and we have

$$E_1 \subset \bigcup_{j \in J_1}^{\infty} Q_j \quad \text{as well as} \quad E_2 \subset \bigcup_{j \in J_2}^{\infty} Q_j.$$

Therefore,

$$\begin{aligned} m_*(E_1) + m_*(E_2) &\leq \sum_{j\in J_1} |Q_j| + \sum_{j\in J_2} |Q_j| \\ &\leq \sum_{j=1}^{\infty} |Q_j| \\ &\leq m_*(E) + \epsilon. \end{aligned}$$

Since ϵ is arbitrary, the proof of Observation 4 is complete.

Observation 5 *If a set E is the countable union of almost disjoint cubes $E = \bigcup_{j=1}^{\infty} Q_j$, then*

$$m_*(E) = \sum_{j=1}^{\infty} |Q_j|.$$

Let $\tilde{Q}_j$ denote a cube strictly contained in Q_j such that $|Q_j| \leq |\tilde{Q}_j| + \epsilon/2^j$, where ϵ is arbitrary but fixed. Then, for every N, the cubes $\tilde{Q}_1, \tilde{Q}_2, \ldots, \tilde{Q}_N$ are disjoint, hence at a finite distance from one another, and repeated applications of Observation 4 imply

$$m_*\left(\bigcup_{j=1}^{N} \tilde{Q}_j\right) = \sum_{j=1}^{N} |\tilde{Q}_j| \geq \sum_{j=1}^{N} \left(|Q_j| - \epsilon/2^j\right).$$

Since $\bigcup_{j=1}^{N} \tilde{Q}_j \subset E$, we conclude that for every integer N,

$$m_*(E) \geq \sum_{j=1}^{N} |Q_j| - \epsilon.$$

In the limit as N tends to infinity we deduce $\sum_{j=1}^{\infty} |Q_j| \leq m_*(E) + \epsilon$ for every $\epsilon > 0$, hence $\sum_{j=1}^{\infty} |Q_j| \leq m_*(E)$. Therefore, combined with Observation 2, our result proves that we have equality.

This last property shows that if a set can be decomposed into almost disjoint cubes, its exterior measure equals the sum of the volumes of the cubes. In particular, by Theorem 1.4 we see that the exterior measure of an open set equals the sum of the volumes of the cubes in a decomposition, and this coincides with our initial guess. Moreover, this also yields a proof that the sum is independent of the decomposition.

One can see from this that the volumes of simple sets that are calculated by elementary calculus agree with their exterior measure. This assertion can be proved most easily once we have developed the requisite tools in integration theory. (See Chapter 2.) In particular, we can then verify that the exterior measure of a ball (either open or closed) equals its volume.

Despite observations 4 and 5, one *cannot* conclude in general that if $E_1 \cup E_2$ is a disjoint union of subsets of $\mathbb{R}^d$, then

$$m_*(E_1 \cup E_2) = m_*(E_1) + m_*(E_2). \tag{3}$$

In fact (3) holds when the sets in question are not highly irregular or "pathological" but are measurable in the sense described below.

3 Measurable sets and the Lebesgue measure

The notion of measurability isolates a collection of subsets in $\mathbb{R}^d$ for which the exterior measure satisfies all our desired properties, including additivity (and in fact countable additivity) for disjoint unions of sets.

There are a number of different ways of defining measurability, but these all turn out to be equivalent. Probably the simplest and most intuitive is the following: A subset E of $\mathbb{R}^d$ is **Lebesgue measurable**, or simply **measurable**, if for any $\epsilon > 0$ there exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and

$$m_*(\mathcal{O} - E) \leq \epsilon.$$

This should be compared to Observation 3, which holds for *all* sets E.

If E is measurable, we define its **Lebesgue measure** (or **measure**) $m(E)$ by

$$m(E) = m_*(E).$$

Clearly, the Lebesgue measure inherits all the features contained in Observations 1 - 5 of the exterior measure.

Immediately from the definition, we find:

Property 1 *Every open set in* $\mathbb{R}^d$ *is measurable.*

Our immediate goal now is to gather various further properties of measurable sets. In particular, we shall prove that the collection of measurable sets behave well under the various operations of set theory: countable unions, countable intersections, and complements.

Property 2 *If $m_*(E) = 0$, then E is measurable. In particular, if F is a subset of a set of exterior measure 0, then F is measurable.*

By Observation 3 of the exterior measure, for every $\epsilon > 0$ there exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and $m_*(\mathcal{O}) \le \epsilon$. Since $(\mathcal{O} - E) \subset \mathcal{O}$, monotonicity implies $m_*(\mathcal{O} - E) \le \epsilon$, as desired.

As a consequence of this property, we deduce that the Cantor set $\mathcal{C}$ in Example 6 is measurable and has measure 0.

Property 3 *A countable union of measurable sets is measurable.*

Suppose $E = \bigcup_{j=1}^\infty E_j$, where each E_j is measurable. Given $\epsilon > 0$, we may choose for each j an open set $\mathcal{O}_j$ with $E_j \subset \mathcal{O}_j$ and $m_*(\mathcal{O}_j - E_j) \le \epsilon/2^j$. Then the union $\mathcal{O} = \bigcup_{j=1}^\infty \mathcal{O}_j$ is open, $E \subset \mathcal{O}$, and $(\mathcal{O} - E) \subset \bigcup_{j=1}^\infty (\mathcal{O}_j - E_j)$, so monotonicity and sub-additivity of the exterior measure imply

$$m_*(\mathcal{O} - E) \le \sum_{j=1}^\infty m_*(\mathcal{O}_j - E_j) \le \epsilon.$$

Property 4 *Closed sets are measurable.*

First, we observe that it suffices to prove that compact sets are measurable. Indeed, any closed set F can be written as the union of compact sets, say $F = \bigcup_{k=1}^\infty F \cap B_k$, where B_k denotes the closed ball of radius k centered at the origin; then Property 3 applies.

So, suppose F is compact (so that in particular $m_*(F) < \infty$), and let $\epsilon > 0$. By Observation 3 we can select an open set $\mathcal{O}$ with $F \subset \mathcal{O}$ and $m_*(\mathcal{O}) \le m_*(F) + \epsilon$. Since F is closed, the difference $\mathcal{O} - F$ is open, and by Theorem 1.4 we may write this difference as a countable union of almost disjoint cubes

$$\mathcal{O} - F = \bigcup_{j=1}^\infty Q_j.$$

For a fixed N, the finite union $K = \bigcup_{j=1}^N Q_j$ is compact; therefore $d(K, F) > 0$ (we isolate this little fact in a lemma below). Since $(K \cup F) \subset \mathcal{O}$, Observations 1, 4, and 5 of the exterior measure imply

$$\begin{aligned} m_*(\mathcal{O}) &\ge m_*(F) + m_*(K) \\ &= m_*(F) + \sum_{j=1}^N m_*(Q_j). \end{aligned}$$

Hence $\sum_{j=1}^N m_*(Q_j) \leq m_*(\mathcal{O}) - m_*(F) \leq \epsilon$, and this also holds in the limit as N tends to infinity. Invoking the sub-additivity property of the exterior measure finally yields

$$m_*(\mathcal{O} - F) \leq \sum_{j=1}^{\infty} m_*(Q_j) \leq \epsilon,$$

as desired.

We digress briefly to complete the above argument by proving the following.

Lemma 3.1 *If F is closed, K is compact, and these sets are disjoint, then $d(F, K) > 0$.*

Proof. Since F is closed, for each point $x \in K$, there exists $\delta_x > 0$ so that $d(x, F) > 3\delta_x$. Since $\bigcup_{x \in K} B_{2\delta_x}(x)$ covers K, and K is compact, we may find a subcover, which we denote by $\bigcup_{j=1}^N B_{2\delta_j}(x_j)$. If we let $\delta = \min(\delta_1, \ldots, \delta_N)$, then we must have $d(K, F) \geq \delta > 0$. Indeed, if $x \in K$ and $y \in F$, then for some j we have $|x_j - x| \leq 2\delta_j$, and by construction $|y - x_j| \geq 3\delta_j$. Therefore

$$|y - x| \geq |y - x_j| - |x_j - x| \geq 3\delta_j - 2\delta_j \geq \delta,$$

and the lemma is proved.

Property 5 *The complement of a measurable set is measurable.*

If E is measurable, then for every positive integer n we may choose an open set $\mathcal{O}_n$ with $E \subset \mathcal{O}_n$ and $m_*(\mathcal{O}_n - E) \leq 1/n$. The complement $\mathcal{O}_n^c$ is closed, hence measurable, which implies that the union $S = \bigcup_{n=1}^{\infty} \mathcal{O}_n^c$ is also measurable by Property 3. Now we simply note that $S \subset E^c$, and

$$(E^c - S) \subset (\mathcal{O}_n - E),$$

such that $m_*(E^c - S) \leq 1/n$ for all n. Therefore, $m_*(E^c - S) = 0$, and $E^c - S$ is measurable by Property 2. Therefore E^c is measurable since it is the union of two measurable sets, namely S and $(E^c - S)$.

Property 6 *A countable intersection of measurable sets is measurable.*

This follows from Properties 3 and 5, since

$$\bigcap_{j=1}^{\infty} E_j = \left(\bigcup_{j=1}^{\infty} E_j^c \right)^c.$$

In conclusion, we find that the family of measurable sets is closed under the familiar operations of set theory. We point out that we have shown more than simply closure with respect to finite unions and intersections: we have proved that the collection of measurable sets is closed under *countable* unions and intersections. This passage from finite operations to infinite ones is crucial in the context of analysis. We emphasize, however, that the operations of *uncountable* unions or intersections are not permissible when dealing with measurable sets!

Theorem 3.2 *If $E_1, E_2, \ldots,$ are disjoint measurable sets, and $E = \bigcup_{j=1}^{\infty} E_j$, then*

$$m(E) = \sum_{j=1}^{\infty} m(E_j).$$

Proof. First, we assume further that each E_j is bounded. Then, for each j, by applying the definition of measurability to E_j^c, we can choose a closed subset F_j of E_j with $m_*(E_j - F_j) \leq \epsilon/2^j$. For each fixed N, the sets $F_1, \ldots, F_N$ are compact and disjoint, so that $m\left(\bigcup_{j=1}^{N} F_j\right) = \sum_{j=1}^{N} m(F_j)$. Since $\bigcup_{j=1}^{N} F_j \subset E$, we must have

$$m(E) \geq \sum_{j=1}^{N} m(F_j) \geq \sum_{j=1}^{N} m(E_j) - \epsilon.$$

Letting N tend to infinity, since ϵ was arbitrary we find that

$$m(E) \geq \sum_{j=1}^{\infty} m(E_j).$$

Since the reverse inequality always holds (by sub-additivity in Observation 2), this concludes the proof when each E_j is bounded.

In the general case, we select any sequence of cubes $\{Q_k\}_{k=1}^{\infty}$ that increases to $\mathbb{R}^d$, in the sense that $Q_k \subset Q_{k+1}$ for all $k \geq 1$ and $\bigcup_{k=1}^{\infty} Q_k = \mathbb{R}^d$. We then let $S_1 = Q_1$ and $S_k = Q_k - Q_{k-1}$ for $k \geq 2$. If we define measurable sets by $E_{j,k} = E_j \cap S_k$, then

$$E = \bigcup_{j,k} E_{j,k}.$$

The union above is disjoint and every $E_{j,k}$ is bounded. Moreover $E_j = \bigcup_{k=1}^{\infty} E_{j,k}$, and this union is also disjoint. Putting these facts together,

and using what has already been proved, we obtain

$$m(E)=\sum_{j,k} m(E_{j,k})=\sum_j\sum_k m(E_{j,k})=\sum_j m(E_j),$$

as claimed.

With this, the countable additivity of the Lebesgue measure on measurable sets has been established. This result provides the necessary connection between the following:

- our primitive notion of volume given by the exterior measure,
- the more refined idea of measurable sets, and
- the countably infinite operations allowed on these sets.

We make two definitions to state succinctly some further consequences.
If $E_1, E_2, \dots$ is a countable collection of subsets of $\mathbb{R}^d$ that increases to E in the sense that $E_k \subset E_{k+1}$ for all k, and $E=\bigcup_{k=1}^\infty E_k$, then we write $E_k \nearrow E$.
Similarly, if $E_1, E_2, \dots$ decreases to E in the sense that $E_k \supset E_{k+1}$ for all k, and $E=\bigcap_{k=1}^\infty E_k$, we write $E_k \searrow E$.

Corollary 3.3 *Suppose $E_1, E_2, \dots$ are measurable subsets of $\mathbb{R}^d$.*

(i) *If $E_k \nearrow E$, then $m(E)=\lim_{N\to\infty} m(E_N)$.*

(ii) *If $E_k \searrow E$ and $m(E_k)<\infty$ for some k, then*

$$m(E)=\lim_{N\to\infty} m(E_N).$$

Proof. For the first part, let $G_1=E_1$, $G_2=E_2-E_1$, and in general $G_k=E_k-E_{k-1}$ for $k\geq 2$. By their construction, the sets G_k are measurable, disjoint, and $E=\bigcup_{k=1}^\infty G_k$. Hence

$$m(E)=\sum_{k=1}^\infty m(G_k)=\lim_{N\to\infty}\sum_{k=1}^N m(G_k)=\lim_{N\to\infty} m\left(\bigcup_{k=1}^N G_k\right),$$

and since $\bigcup_{k=1}^N G_k=E_N$ we get the desired limit.
For the second part, we may clearly assume that $m(E_1)<\infty$. Let $G_k=E_k-E_{k+1}$ for each k, so that

$$E_1=E\cup\bigcup_{k=1}^\infty G_k$$

is a disjoint union of measurable sets. As a result, we find that

$$\begin{aligned} m(E_1) &= m(E) + \lim_{N\to\infty} \sum_{k=1}^{N-1} (m(E_k) - m(E_{k+1})) \\ &= m(E) + m(E_1) - \lim_{N\to\infty} m(E_N). \end{aligned}$$

Hence, since $m(E_1) < \infty$, we see that $m(E) = \lim_{N\to\infty} m(E_N)$, and the proof is complete.

The reader should note that the second conclusion may fail without the assumption that $m(E_k) < \infty$ for some k. This is shown by the simple example when $E_n = (n, \infty) \subset \mathbb{R}$, for all n.

What follows provides an important geometric and analytic insight into the nature of measurable sets, in terms of their relation to open and closed sets. Its thrust is that, in effect, an arbitrary measurable set can be well approximated by the open sets that contain it, and alternatively, by the closed sets it contains.

Theorem 3.4 *Suppose E is a measurable subset of $\mathbb{R}^d$. Then, for every $\epsilon > 0$:*

(i) *There exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and $m(\mathcal{O} - E) \le \epsilon$.*

(ii) *There exists a closed set F with $F \subset E$ and $m(E - F) \le \epsilon$.*

(iii) *If $m(E)$ is finite, there exists a compact set K with $K \subset E$ and $m(E - K) \le \epsilon$.*

(iv) *If $m(E)$ is finite, there exists a finite union $F = \bigcup_{j=1}^{N} Q_j$ of closed cubes such that*

$$m(E \triangle F) \le \epsilon.$$

The notation $E \triangle F$ stands for the **symmetric difference** between the sets E and F, defined by $E \triangle F = (E - F) \cup (F - E)$, which consists of those points that belong to only one of the two sets E or F.

Proof. Part (i) is just the definition of measurability. For the second part, we know that E^c is measurable, so there exists an open set $\mathcal{O}$ with $E^c \subset \mathcal{O}$ and $m(\mathcal{O} - E^c) \le \epsilon$. If we let $F = \mathcal{O}^c$, then F is closed, $F \subset E$, and $E - F = \mathcal{O} - E^c$. Hence $m(E - F) \le \epsilon$ as desired.

For (iii), we first pick a closed set F so that $F \subset E$ and $m(E - F) \le \epsilon/2$. For each n, we let B_n denote the ball centered at the origin of radius

n, and define compact sets $K_n = F \cap B_n$. Then $E - K_n$ is a sequence of measurable sets that decreases to $E - F$, and since $m(E) < \infty$, we conclude that for all large n one has $m(E - K_n) \leq \epsilon$.

For the last part, choose a family of closed cubes $\{Q_j\}_{j=1}^{\infty}$ so that

$$E \subset \bigcup_{j=1}^{\infty} Q_j \quad \text{and} \quad \sum_{j=1}^{\infty} |Q_j| \leq m(E) + \epsilon/2.$$

Since $m(E) < \infty$, the series converges and there exists $N > 0$ such that $\sum_{j=N+1}^{\infty} |Q_j| < \epsilon/2$. If $F = \bigcup_{j=1}^{N} Q_j$, then

$$\begin{aligned} m(E \triangle F) &= m(E - F) + m(F - E) \\ &\leq m\left(\bigcup_{j=N+1}^{\infty} Q_j\right) + m\left(\bigcup_{j=1}^{\infty} Q_j - E\right) \\ &\leq \sum_{j=N+1}^{\infty} |Q_j| + \sum_{j=1}^{\infty} |Q_j| - m(E) \\ &\leq \epsilon. \end{aligned}$$

Invariance properties of Lebesgue measure

A crucial property of Lebesgue measure in $\mathbb{R}^d$ is its translation-invariance, which can be stated as follows: if E is a measurable set and $h \in \mathbb{R}^d$, then the set $E_h = E + h = \{x + h : x \in E\}$ is also measurable, and $m(E + h) = m(E)$. With the observation that this holds for the special case when E is a cube, one passes to the exterior measure of arbitrary sets E, and sees from the definition of m_* given in Section 2 that $m_*(E_h) = m_*(E)$. To prove the measurability of E_h under the assumption that E is measurable, we note that if $\mathcal{O}$ is open, $\mathcal{O} \supset E$, and $m_*(\mathcal{O} - E) < \epsilon$, then $\mathcal{O}_h$ is open, $\mathcal{O}_h \supset E_h$, and $m_*(\mathcal{O}_h - E_h) < \epsilon$.

In the same way one can prove the relative dilation-invariance of Lebesgue measure. Suppose $\delta > 0$, and denote by δE the set $\{\delta x : x \in E\}$. We can then assert that δE is measurable whenever E is, and $m(\delta E) = \delta^d m(E)$. One can also easily see that Lebesgue measure is reflection-invariant. That is, whenever E is measurable, so is $-E = \{-x : x \in E\}$ and $m(-E) = m(E)$.

Other invariance properties of Lebesgue measure are in Exercise 7 and 8, and Problem 4 of Chapter 2.

σ-algebras and Borel sets

A **σ-algebra** of sets is a collection of subsets of $\mathbb{R}^d$ that is closed under countable unions, countable intersections, and complements.

The collection of all subsets of $\mathbb{R}^d$ is of course a σ-algebra. A more interesting and relevant example consists of all measurable sets in $\mathbb{R}^d$, which we have just shown also forms a σ-algebra.

Another σ-algebra, which plays a vital role in analysis, is the **Borel σ-algebra** in $\mathbb{R}^d$, denoted by $\mathcal{B}_{\mathbb{R}^d}$, which by definition is the smallest σ-algebra that contains all open sets. Elements of this σ-algebra are called **Borel sets**.

The definition of the Borel σ-algebra will be meaningful once we have defined the term "smallest," and shown that such a σ-algebra exists and is unique. The term "smallest" means that if $\mathcal{S}$ is any σ-algebra that contains all open sets in $\mathbb{R}^d$, then necessarily $\mathcal{B}_{\mathbb{R}^d} \subset \mathcal{S}$. Since we observe that any intersection (not necessarily countable) of σ-algebras is again a σ-algebra, we may define $\mathcal{B}_{\mathbb{R}^d}$ as the intersection of all σ-algebras that contain the open sets. This shows the existence and uniqueness of the Borel σ-algebra.

Since open sets are measurable, we conclude that the Borel σ-algebra is contained in the σ-algebra of measurable sets. Naturally, we may ask if this inclusion is strict: do there exist Lebesgue measurable sets which are not Borel sets? The answer is "yes." (See Exercise 35.)

From the point of view of the Borel sets, the Lebesgue sets arise as the **completion** of the σ-algebra of Borel sets, that is, by adjoining all subsets of Borel sets of measure zero. This is an immediate consequence of Corollary 3.5 below.

Starting with the open and closed sets, which are the simplest Borel sets, one could try to list the Borel sets in order of their complexity. Next in order would come countable intersections of open sets; such sets are called $\boldsymbol{G_\delta}$ sets. Alternatively, one could consider their complements, the countable union of closed sets, called the $\boldsymbol{F_\sigma}$ sets.[3]

Corollary 3.5 *A subset E of $\mathbb{R}^d$ is measurable*

(i) *if and only if E differs from a G_δ by a set of measure zero,*

(ii) *if and only if E differs from an F_σ by a set of measure zero.*

Proof. Clearly E is measurable whenever it satisfies either (i) or (ii), since the F_σ, G_δ, and sets of measure zero are measurable.

---

[3]The terminology G_δ comes from German "Gebiete" and "Durschnitt"; F_σ comes from French "fermé" and "somme."

Conversely, if E is measurable, then for each integer $n \geq 1$ we may select an open set $\mathcal{O}_n$ that contains E, and such that $m(\mathcal{O}_n - E) \leq 1/n$. Then $S = \bigcap_{n=1}^\infty \mathcal{O}_n$ is a G_δ that contains E, and $(S - E) \subset (\mathcal{O}_n - E)$ for all n. Therefore $m(S - E) \leq 1/n$ for all n; hence $S - E$ has exterior measure zero, and is therefore measurable.

For the second implication, we simply apply part (ii) of Theorem 3.4 with $\epsilon = 1/n$, and take the union of the resulting closed sets.

Construction of a non-measurable set

Are all subsets of $\mathbb{R}^d$ measurable? In this section, we answer this question when $d = 1$ by constructing a subset of $\mathbb{R}$ which is *not* measurable.[4] This justifies the conclusion that a satisfactory theory of measure cannot encompass *all* subsets of $\mathbb{R}$.

The construction of a non-measurable set $\mathcal{N}$ uses the axiom of choice, and rests on a simple equivalence relation among real numbers in $[0, 1]$.

We write $x \sim y$ whenever $x - y$ is rational, and note that this is an equivalence relation since the following properties hold:

- $x \sim x$ for every $x \in [0, 1]$
- if $x \sim y$, then $y \sim x$
- if $x \sim y$ and $y \sim z$, then $x \sim z$.

Two equivalence classes either are disjoint or coincide, and $[0, 1]$ is the disjoint union of all equivalence classes, which we write as

$$[0, 1] = \bigcup_\alpha \mathcal{E}_\alpha.$$

Now we construct the set $\mathcal{N}$ by choosing exactly one element x_α from each $\mathcal{E}_\alpha$, and setting $\mathcal{N} = \{x_\alpha\}$. This (seemingly obvious) step requires further comment, which we postpone until after the proof of the following theorem.

Theorem 3.6 *The set $\mathcal{N}$ is not measurable.*

The proof is by contradiction, so we assume that $\mathcal{N}$ is measurable. Let $\{r_k\}_{k=1}^\infty$ be an enumeration of all the rationals in $[-1, 1]$, and consider the translates

$$\mathcal{N}_k = \mathcal{N} + r_k.$$

[4]The existence of such a set in $\mathbb{R}$ implies the existence of corresponding non-measurable subsets of $\mathbb{R}^d$ for each d, as a consequence of Proposition 3.4 in the next chapter.

We claim that the sets $\mathcal{N}_k$ are disjoint, and

$$[0,1] \subset \bigcup_{k=1}^{\infty} \mathcal{N}_k \subset [-1,2]. \tag{4}$$

To see why these sets are disjoint, suppose that the intersection $\mathcal{N}_k \cap \mathcal{N}_{k'}$ is non-empty. Then there exist rationals $r_k \neq r'_k$ and α and β with $x_\alpha + r_k = x_\beta + r_{k'}$; hence

$$x_\alpha - x_\beta = r_{k'} - r_k.$$

Consequently $\alpha \neq \beta$ and $x_\alpha - x_\beta$ is rational; hence $x_\alpha \sim x_\beta$, which contradicts the fact that $\mathcal{N}$ contains only *one* representative of each equivalence class.

The second inclusion is straightforward since each $\mathcal{N}_k$ is contained in $[-1,2]$ by construction. Finally, if $x \in [0,1]$, then $x \sim x_\alpha$ for some α, and therefore $x - x_\alpha = r_k$ for some k. Hence $x \in \mathcal{N}_k$, and the first inclusion holds.

Now we may conclude the proof of the theorem. If $\mathcal{N}$ were measurable, then so would be $\mathcal{N}_k$ for all k, and since the union $\bigcup_{k=1}^{\infty} \mathcal{N}_k$ is disjoint, the inclusions in (4) yield

$$1 \leq \sum_{k=1}^{\infty} m(\mathcal{N}_k) \leq 3.$$

Since $\mathcal{N}_k$ is a translate of $\mathcal{N}$, we must have $m(\mathcal{N}_k) = m(\mathcal{N})$ for all k. Consequently,

$$1 \leq \sum_{k=1}^{\infty} m(\mathcal{N}) \leq 3.$$

This is the desired contradiction, since neither $m(\mathcal{N}) = 0$ nor $m(\mathcal{N}) > 0$ is possible.

Axiom of choice

That the construction of the set $\mathcal{N}$ is possible is based on the following general proposition.

- Suppose E is a set and $\{E_\alpha\}$ is a collection of non-empty subsets of E. (The indexing set of α's is not assumed to be countable.) Then there is a function $\alpha \mapsto x_\alpha$ (a "choice function") such that $x_\alpha \in E_\alpha$, for all α.

In this general form this assertion is known as the **axiom of choice**. This axiom occurs (at least implicitly) in many proofs in mathematics, but because of its seeming intuitive self-evidence, its significance was not at first understood. The initial realization of the importance of this axiom was in its use to prove a famous assertion of Cantor, the **well-ordering principle**. This proposition (sometimes referred to as "transfinite induction") can be formulated as follows.

A set E is **linearly ordered** if there is a binary relation $\leq$ such that:

(a) $x \leq x$ for all $x \in E$.

(b) If $x, y \in E$ are distinct, then either $x \leq y$ or $y \leq x$ (but not both).

(c) If $x \leq y$ and $y \leq z$, then $x \leq z$.

We say that a set E can be **well-ordered** if it can be linearly ordered in such a way that *every* non-empty subset $A \subset E$ has a smallest element in that ordering (that is, an element $x_0 \in A$ such that $x_0 \leq x$ for any other $x \in A$).

A simple example of a well-ordered set is $\mathbb{Z}^+$, the positive integers with their usual ordering. The fact that $\mathbb{Z}^+$ is well-ordered is an essential part of the usual (finite) induction principle. More generally, the well-ordering principle states:

- Any set E can be well-ordered.

It is in fact nearly obvious that the well-ordering principle implies the axiom of choice: if we well-order E, we can choose x_α to be the smallest element in E_α, and in this way we have constructed the required choice function. It is also true, but not as easy to show, that the converse implication holds, namely that the axiom of choice implies the well-ordering principle. (See Problem 6 for another equivalent formulation of the Axiom of Choice.)

We shall follow the common practice of assuming the axiom of choice (and hence the validity of the well-ordering principle).[5] However, we should point out that while the axiom of choice seems self-evident the well-ordering principle leads quickly to some baffling conclusions: one only needs to spend a little time trying to imagine what a well-ordering of the reals might look like!

[5]It can be proved that in an appropriate formulation of the axioms of set theory, the axiom of choice is independent of the other axioms; thus we are free to accept its validity.

4 Measurable functions

With the notion of measurable sets in hand, we now turn our attention to the objects that lie at the heart of integration theory: measurable functions.

The starting point is the notion of a **characteristic function** of a set E, which is defined by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{if } x \notin E. \end{cases}$$

The next step is to pass to the functions that are the building blocks of integration theory. For the Riemann integral it is in effect the class of **step functions**, with each given as a finite sum

$$f = \sum_{k=1}^{N} a_k \chi_{R_k}, \tag{5}$$

where each R_k is a rectangle, and the a_k are constants.

However, for the Lebesgue integral we need a more general notion, as we shall see in the next chapter. A **simple function** is a finite sum

$$f = \sum_{k=1}^{N} a_k \chi_{E_k} \tag{6}$$

where each E_k is a measurable set of finite measure, and the a_k are constants.

4.1 Definition and basic properties

We begin by considering only real-valued functions f on $\mathbb{R}^d$, which we allow to take on the infinite values $+\infty$ and $-\infty$, so that $f(x)$ belongs to the extended real numbers

$$-\infty \le f(x) \le \infty.$$

We shall say that f is **finite-valued** if $-\infty < f(x) < \infty$ for all x. In the theory that follows, and the many applications of it, we shall almost always find ourselves in situations where a function takes on infinite values on at most a set of measure zero.

A function f defined on a measurable subset E of $\mathbb{R}^d$ is **measurable**, if for all $a \in \mathbb{R}$, the set

$$f^{-1}([-\infty, a)) = \{x \in E : f(x) < a\}$$

is measurable. To simplify our notation, we shall often denote the set $\{x \in E : f(x) < a\}$ simply by $\{f < a\}$ whenever no confusion is possible.

First, we note that there are many equivalent definitions of measurable functions. For example, we may require instead that the inverse image of closed intervals be measurable. Indeed, to prove that f is measurable if and only if $\{x : f(x) \leq a\} = \{f \leq a\}$ is measurable for every a, we note that in one direction, one has

$$\{f \leq a\} = \bigcap_{k=1}^{\infty} \{f < a + 1/k\},$$

and recall that the countable intersection of measurable sets is measurable. For the other direction, we observe that

$$\{f < a\} = \bigcup_{k=1}^{\infty} \{f \leq a - 1/k\}.$$

Similarly, f is measurable if and only if $\{f \geq a\}$ (or $\{f > a\}$) is measurable for every a. In the first case this is immediate from our definition and the fact that $\{f \geq a\}$ is the complement of $\{f < a\}$, and in the second case this follows from what we have just proved and the fact that $\{f \leq a\} = \{f > a\}^c$. A simple consequence is that $-f$ is measurable whenever f is measurable.

In the same way, one can show that if f is finite-valued, then it is measurable if and only if the sets $\{a < f < b\}$ are measurable for every $a, b \in \mathbb{R}$. Similar conclusions hold for whichever combination of strict or weak inequalities one chooses. For example, if f is finite-valued, then it is measurable if and only if $\{a \leq f < b\}$ for all $a, b \in \mathbb{R}$. By the same arguments one sees the following:

Property 1 *The finite-valued function f is measurable if and only if $f^{-1}(\mathcal{O})$ is measurable for every open set $\mathcal{O}$, and if and only if $f^{-1}(F)$ is measurable for every closed set F.*

Note that this property also applies to extended-valued functions, if we make the additional hypothesis that both $f^{-1}(\infty)$ and $f^{-1}(-\infty)$ are measurable sets.

Property 2 *If f is continuous on $\mathbb{R}^d$, then f is measurable. If f is measurable and finite-valued, and Φ is continuous, then $\Phi \circ f$ is measurable.*

In fact, Φ is continuous, so $\Phi^{-1}((-\infty, a))$ is an open set $\mathcal{O}$, and hence $(\Phi \circ f)^{-1}((-\infty, a)) = f^{-1}(\mathcal{O})$ is measurable.

It should be noted, however, that in general it is not true that $f \circ \Phi$ is measurable whenever f is measurable and Φ is continuous. See Exercise 35.

Property 3 *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions. Then*

$$\sup_n f_n(x), \quad \inf_n f_n(x), \quad \limsup_{n\to\infty}, f_n(x) \quad \textit{and} \quad \liminf_{n\to\infty} f_n(x)$$

are measurable.

Proving that $\sup_n f_n$ is measurable requires noting that $\{\sup_n f_n > a\} = \bigcup_n \{f_n > a\}$. This also yields the result for $\inf_n f_n(x)$, since this quantity equals $-\sup_n(-f_n(x))$.

The result for the limsup and liminf also follows from the two observations

$$\limsup_{n\to\infty} f_n(x) = \inf_k \{\sup_{n\geq k} f_n\} \quad \text{and} \quad \liminf_{n\to\infty} f_n(x) = \sup_k \{\inf_{n\geq k} f_n\}.$$

Property 4 *If $\{f_n\}_{n=1}^{\infty}$ is a collection of measurable functions, and*

$$\lim_{n\to\infty} f_n(x) = f(x),$$

then f is measurable.

Since $f(x) = \limsup_{n\to\infty} f_n(x) = \liminf_{n\to\infty} f_n(x)$, this property is a consequence of property 3.

Property 5 *If f and g are measurable, then*

(i) *The integer powers f^k, $k \geq 1$ are measurable.*

(ii) *$f + g$ and fg are measurable if both f and g are finite-valued.*

For (i) we simply note that if k is odd, then $\{f^k > a\} = \{f > a^{1/k}\}$, and if k is even and $a \geq 0$, then $\{f^k > a\} = \{f > a^{1/k}\} \cup \{f < -a^{1/k}\}$.

For (ii), we first see that $f + g$ is measurable because

$$\{f + g > a\} = \bigcup_{r\in\mathbb{Q}} \{f > a - r\} \cap \{g > r\},$$

with $\mathbb{Q}$ denoting the rationals.

Finally, fg is measurable because of the previous results and the fact that

$$fg = \frac{1}{4}[(f+g)^2 - (f-g)^2].$$

We shall say that two functions f and g defined on a set E are equal **almost everywhere**, and write

$$f(x) = g(x) \qquad \text{a.e. } x \in E,$$

if the set $\{x \in E : f(x) \neq g(x)\}$ has measure zero. We sometimes abbreviate this by saying that $f = g$ a.e. More generally, a property or statement is said to hold almost everywhere (a.e.) if it is true except on a set of measure zero.

One sees easily that if f is measurable and $f = g$ a.e., then g is measurable. This follows at once from the fact that $\{f < a\}$ and $\{g < a\}$ differ by a set of measure zero. Moreover, all the properties above can be relaxed to conditions holding almost everywhere. For instance, if $\{f_n\}_{n=1}^{\infty}$ is a collection of measurable functions, and

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \text{a.e.},$$

then f is measurable.

Note that if f and g are defined almost everywhere on a measurable subset $E \subset \mathbb{R}^d$, then the functions $f + g$ and fg can only be defined on the intersection of the domains of f and g. Since the union of two sets of measure zero has again measure zero, $f + g$ is defined almost everywhere on E. We summarize this discussion as follows.

Property 6 *Suppose f is measurable, and $f(x) = g(x)$ for a.e. x. Then g is measurable.*

In this light, Property 5 (ii) also holds when f and g are finite-valued almost everywhere.

4.2 Approximation by simple functions or step functions

The theorems in this section are all of the same nature and provide further insight in the structure of measurable functions. We begin by approximating pointwise, non-negative measurable functions by simple functions.

Theorem 4.1 *Suppose f is a non-negative measurable function on $\mathbb{R}^d$. Then there exists an increasing sequence of non-negative simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that converges pointwise to f, namely,*

$$\varphi_k(x) \leq \varphi_{k+1}(x) \quad \textit{and} \quad \lim_{k\to\infty} \varphi_k(x) = f(x), \textit{ for all } x.$$

Proof. We begin first with a truncation. For $N \geq 1$, let Q_N denote the cube centered at the origin and of side length N. Then we define

$$F_N(x) = \begin{cases} f(x) & \text{if } x \in Q_N \text{ and } f(x) \leq N, \\ N & \text{if } x \in Q_N \text{ and } f(x) > N, \\ 0 & \text{otherwise.} \end{cases}$$

Then, $F_N(x) \to f(x)$ as N tends to infinity for all x. Now, we partition the range of F_N, namely $[0, N]$, as follows. For fixed $N, M \geq 1$, we define

$$E_{\ell,M} = \left\{ x \in Q_N : \frac{\ell}{M} < F_N(x) \leq \frac{\ell+1}{M} \right\}, \qquad \text{for } 0 \leq \ell < NM.$$

Then we may form

$$F_{N,M}(x) = \sum_{\ell} \frac{\ell}{M} \, \chi_{E_{\ell,M}}(x).$$

Each $F_{N,M}$ is a simple function that satisfies $0 \leq F_N(x) - F_{N,M}(x) \leq 1/M$ for all x. If we now choose $N = M = 2^k$ with $k \geq 1$ integral, and let $\varphi_k = F_{2^k,2^k}$, then we see that $0 \leq F_M(x) - \varphi_k(x) \leq 1/2^k$ for all x, $\{\varphi_k\}$ is increasing, and this sequence satisfies all the desired properties.

Note that the result holds for non-negative functions that are extended-valued, if the limit $+\infty$ is allowed. We now drop the assumption that f is non-negative, and also allow the extended limit $-\infty$.

Theorem 4.2 *Suppose f is measurable on $\mathbb{R}^d$. Then there exists a sequence of simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that satisfies*

$$|\varphi_k(x)| \leq |\varphi_{k+1}(x)| \quad \textit{and} \quad \lim_{k\to\infty} \varphi_k(x) = f(x), \textit{ for all } x.$$

In particular, we have $|\varphi_k(x)| \leq |f(x)|$ for all x and k.

Proof. We use the following decomposition of the function f: $f(x) = f^+(x) - f^-(x)$, where

$$f^+(x) = \max(f(x), 0) \quad \text{and} \quad f^-(x) = \max(-f(x), 0).$$

Since both f^+ and f^- are non-negative, the previous theorem yields two increasing sequences of non-negative simple functions $\{\varphi_k^{(1)}(x)\}_{k=1}^\infty$ and $\{\varphi_k^{(2)}(x)\}_{k=1}^\infty$ which converge pointwise to f^+ and f^-, respectively. Then, if we let

$$\varphi_k(x) = \varphi_k^{(1)}(x) - \varphi_k^{(2)}(x),$$

we see that $\varphi_k(x)$ converges to $f(x)$ for all x. Finally, the sequence $\{|\varphi_k|\}$ is increasing because the definition of f^+, f^- and the properties of $\varphi_k^{(1)}$ and $\varphi_k^{(2)}$ imply that

$$|\varphi_k(x)| = \varphi_k^{(1)}(x) + \varphi_k^{(2)}(x).$$

We may now go one step further, and approximate by step functions. Here, in general, the convergence may hold only almost everywhere.

Theorem 4.3 *Suppose f is measurable on $\mathbb{R}^d$. Then there exists a sequence of step functions $\{\psi_k\}_{k=1}^\infty$ that converges pointwise to $f(x)$ for almost every x.*

Proof. By the previous result, it suffices to show that if E is a measurable set with finite measure, then $f = \chi_E$ can be approximated by step functions. To this end, we recall part (iv) of Theorem 3.4, which states that for every ϵ there exist cubes $Q_1, \ldots, Q_N$ such that $m(E \triangle \bigcup_{j=1}^N Q_j) \leq \epsilon$. By considering the grid formed by extending the sides of these cubes, we see that there exist almost disjoint rectangles $\tilde{R}_1, \ldots, \tilde{R}_M$ such that $\bigcup_{j=1}^N Q_j = \bigcup_{j=1}^M \tilde{R}_j$. By taking rectangles R_j contained in $\tilde{R}_j$, and slightly smaller in size, we find a collection of disjoint rectangles that satisfy $m(E \triangle \bigcup_{j=1}^M R_j) \leq 2\epsilon$. Therefore

$$f(x) = \sum_{j=1}^M \chi_{R_j}(x),$$

except possibly on a set of measure $\leq 2\epsilon$. Consequently, for every $k \geq 1$, there exists a step function $\psi_k(x)$ such that if

$$E_k = \{x : \ f(x) \neq \psi_k(x)\},$$

then $m(E_k) \leq 2^{-k}$. If we let $F_K = \bigcup_{j=K+1}^\infty E_j$ and $F = \bigcap_{K=1}^\infty F_K$, then $m(F) = 0$ since $m(F_K) \leq 2^{-K}$, and $\psi_k(x) \to f(x)$ for all x in the complement of F, which is the desired result.

4.3 Littlewood's three principles

Although the notions of measurable sets and measurable functions represent new tools, we should not overlook their relation to the older concepts they replaced. Littlewood aptly summarized these connections in the form of three principles that provide a useful intuitive guide in the initial study of the theory.

(i) Every set is nearly a finite union of intervals.

(ii) Every function is nearly continuous.

(iii) Every convergent sequence is nearly uniformly convergent.

The sets and functions referred to above are of course assumed to be measurable. The catch is in the word "nearly," which has to be understood appropriately in each context. A precise version of the first principle appears in part (iv) of Theorem 3.4. An exact formulation of the third principle is given in the following important result.

Theorem 4.4 (Egorov) *Suppose* $\{f_k\}_{k=1}^{\infty}$ *is a sequence of measurable functions defined on a measurable set* E *with* $m(E) < \infty$*, and assume that* $f_k \to f$ *a.e on* E*. Given* $\epsilon > 0$*, we can find a closed set* $A_\epsilon \subset E$ *such that* $m(E - A_\epsilon) \leq \epsilon$ *and* $f_k \to f$ *uniformly on* A_ϵ*.*

Proof. We may assume without loss of generality that $f_k(x) \to f(x)$ for every $x \in E$. For each pair of non-negative integers n and k, let

$$E_k^n = \{x \in E : |f_j(x) - f(x)| < 1/n, \quad \text{for all } j > k\}.$$

Now fix n and note that $E_k^n \subset E_{k+1}^n$, and $E_k^n \nearrow E$ as k tends to infinity. By Corollary 3.3, we find that there exists k_n such that $m(E - E_{k_n}^n) < 1/2^n$. By construction, we then have

$$|f_j(x) - f(x)| < 1/n \qquad \text{whenever } j > k_n \text{ and } x \in E_{k_n}^n.$$

We choose N so that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon/2$, and let

$$\tilde{A}_\epsilon = \bigcap_{n \geq N} E_{k_n}^n.$$

We first observe that

$$m(E - \tilde{A}_\epsilon) \leq \sum_{n=N}^{\infty} m(E - E_{k_n}^n) < \epsilon/2.$$

Next, if $\delta > 0$, we choose $n \geq N$ such that $1/n < \delta$, and note that $x \in \tilde{A}_\epsilon$ implies $x \in E^n_{k_n}$. We see therefore that $|f_j(x) - f(x)| < \delta$ whenever $j > k_n$. Hence f_k converges uniformly to f on $\tilde{A}_\epsilon$.

Finally, using Theorem 3.4 choose a closed subset $A_\epsilon \subset \tilde{A}_\epsilon$ with $m(\tilde{A}_\epsilon - A_\epsilon) < \epsilon/2$. As a result, we have $m(E - A_\epsilon) < \epsilon$ and the theorem is proved.

The next theorem attests to the validity of the second of Littlewood's principle.

Theorem 4.5 (Lusin) *Suppose f is measurable and finite valued on E with E of finite measure. Then for every $\epsilon > 0$ there exists a closed set F_ϵ, with*

$$F_\epsilon \subset E, \quad \textit{and} \quad m(E - F_\epsilon) \leq \epsilon$$

and such that $f|_{F_\epsilon}$ is continuous.

By $f|_{F_\epsilon}$ we mean the restriction of f to the set F_ϵ. The conclusion of the theorem states that if f is viewed as a function defined only on F_ϵ, then f is continuous. However, the theorem does *not* make the stronger assertion that the function f defined on E is continuous at the points of F_ϵ.

Proof. Let f_n be a sequence of step functions so that $f_n \to f$ a.e. Then we may find sets E_n so that $m(E_n) < 1/2^n$ and f_n is continuous outside E_n. By Egorov's theorem, we may find a set $A_{\epsilon/3}$ on which $f_n \to f$ uniformly and $m(E - A_{\epsilon/3}) \leq \epsilon/3$. Then we consider

$$F' = A_{\epsilon/3} - \bigcup_{n \geq N} E_n$$

for N so large that $\sum_{n \geq N} 1/2^n < \epsilon/3$. Now for every $n \geq N$ the function f_n is continuous on F'; thus f (being the uniform limit of $\{f_n\}$) is also continuous on F'. To finish the proof, we merely need to approximate the set F' by a closed set $F_\epsilon \subset F'$ such that $m(F' - F_\epsilon) < \epsilon/3$.

5* The Brunn-Minkowski inequality

Since addition and multiplication by scalars are basic features of vector spaces, it is not surprising that properties of these operations arise in a fundamental way in the theory of Lebesgue measure on $\mathbb{R}^d$. We have already discussed in this connection the translation-invariance and relative

dilation-invariance of Lebesgue measure. Here we come to the study of the sum of two measurable sets A and B, defined as

$$A + B = \{x \in \mathbb{R}^d : \ x = x' + x'' \ \text{ with } x' \in A \text{ and } x'' \in B\}.$$

This notion is of importance in a number of questions, in particular in the theory of convex sets; we shall apply it to the isoperimetric problem in Chapter 3.

In this regard the first (admittedly vague) question we can pose is whether one can establish any general estimate for the measure of $A + B$ in terms of the measures of A and B (assuming that these three sets are measurable). We can see easily that it is not possible to obtain an upper bound for $m(A + B)$ in terms of $m(A)$ and $m(B)$. Indeed, simple examples show that we may have $m(A) = m(B) = 0$ while $m(A + B) > 0$. (See Exercise 20.)

In the converse direction one might ask for a general estimate of the form

$$m(A + B)^\alpha \geq c_\alpha \left(m(A)^\alpha + m(B)^\alpha\right),$$

where α is a positive number and the constant c_α is independent of A and B. Clearly, the best one can hope for is $c_\alpha = 1$. The role of the exponent α can be understood by considering **convex sets**. Such sets A are defined by the property that whenever x and y are in A then the line segment joining them, $\{xt + y(1 - t) : \ 0 \leq t \leq 1\}$, also belongs to A. If we recall the definition $\lambda A = \{\lambda x, \ x \in A\}$ for $\lambda > 0$, we note that whenever A is convex, then $A + \lambda A = (1 + \lambda)A$. However, $m((1 + \lambda)A) = (1 + \lambda)^d m(A)$, and thus the presumed inequality can hold only if $(1 + \lambda)^{d\alpha} \geq 1 + \lambda^{d\alpha}$, for all $\lambda > 0$. Now

$$(a + b)^\gamma \geq a^\gamma + b^\gamma \qquad \text{if } \gamma \geq 1 \text{ and } a, b \geq 0, \tag{7}$$

while the reverse inequality holds if $0 \leq \gamma \leq 1$. (See Exercise 38.) This yields $\alpha \geq 1/d$. Moreover, (7) shows that the inequality with the exponent $1/d$ implies the corresponding inequality with $\alpha \geq 1/d$, and so we are naturally led to the inequality

$$m(A + B)^{1/d} \geq m(A)^{1/d} + m(B)^{1/d}. \tag{8}$$

Before proceeding with the proof of (8), we need to mention a technical impediment that arises. While we may assume that A and B are measurable, it does not necessarily follow that then $A + B$ is measurable. (See Exercise 13 in the next chapter.) However it is easily seen that this

difficulty does not occur when, for example, A and B are closed sets, or when one of them is open. (See Exercise 19.)

With the above considerations in mind we can state the main result.

Theorem 5.1 *Suppose A and B are measurable sets in $\mathbb{R}^d$ and their sum $A+B$ is also measurable. Then the inequality* (8) *holds.*

Let us first check (8) when A and B are rectangles with side lengths $\{a_j\}_{j=1}^d$ and $\{b_j\}_{j=1}^d$, respectively. Then (8) becomes

$$\left(\prod_{j=1}^{d}(a_j+b_j)\right)^{1/d} \geq \left(\prod_{j=1}^{d}a_j\right)^{1/d} + \left(\prod_{j=1}^{d}b_j\right)^{1/d}, \tag{9}$$

which by homogeneity we can reduce to the special case where $a_j + b_j = 1$ for each j. In fact, notice that if we replace a_j, b_j by $\lambda_j a_j, \lambda_j b_j$, with $\lambda_j > 0$, then both sides of (9) are multiplied by $(\lambda_1\lambda_2\cdots\lambda_d)^{1/d}$. We then need only choose $\lambda_j = (a_j+b_j)^{-1}$. With this reduction, the inequality (9) is an immediate consequence of the arithmetic-geometric inequality (Exercise 39)

$$\frac{1}{d}\sum_{j=1}^{d}x_j \geq \left(\prod_{j=1}^{d}x_j\right)^{1/d}, \qquad \text{for all } x_j \geq 0:$$

we add the two inequalities that result when we set $x_j = a_j$ and $x_j = b_j$, respectively.

We next turn to the case when each A and B are the union of finitely many rectangles whose interiors are disjoint. We shall prove (8) in this case by induction on the total number of rectangles in A and B. We denote this number by n. Here it is important to note that the desired inequality is unchanged when we translate A and B independently. In fact, replacing A by $A+h$ and B by $B+h'$ replaces $A+B$ by $A+B+h+h'$, and thus the corresponding measures remain the same. We now choose a pair of disjoint rectangles R_1 and R_2 in the collection making up A, and we note that they can be separated by a coordinate hyperplane. Thus we may assume that for some j, after translation by an appropriate h, R_1 lies in $A_- = A\cap\{x_j \leq 0\}$, and R_2 in $A_+ = A\cap\{0\leq x_j\}$. Observe also that both A_+ and A_- contain at least one less rectangle than A does, and $A = A_-\cup A_+$.

We next translate B so that $B_- = B\cap\{x_j\leq 0\}$ and $B_+ = B\cap\{x_j \geq 0\}$ satisfy

$$\frac{m(B_\pm)}{m(B)} = \frac{m(A_\pm)}{m(A)}.$$

However, $A + B \supset (A_+ + B_+) \cup (A_- + B_-)$, and the union on the right is essentially disjoint, since the two parts lie in different half-spaces. Moreover, the total number of rectangles in either A_+ and B_+, or A_- and B_- is also less than n. Thus the induction hypothesis applies and

$$\begin{aligned} m(A+B) &\geq m(A_+ + B_+) + m(A_- + B_-) \\ &\geq \left(m(A_+)^{1/d} + m(B_+)^{1/d}\right)^d + \left(m(A_-)^{1/d} + m(B_-)^{1/d}\right)^d \\ &= m(A_+)\left[1 + \left(\frac{m(B)}{m(A)}\right)^{1/d}\right]^d + m(A_-)\left[1 + \left(\frac{m(B)}{m(A)}\right)^{1/d}\right]^d \\ &= \left(m(A)^{1/d} + m(B)^{1/d}\right)^d, \end{aligned}$$

which gives the desired inequality (8) when A and B are both finite unions of rectangles with disjoint interiors.

Next, this quickly implies the result when A and B are open sets of finite measure. Indeed, by Theorem 1.4, for any $\epsilon > 0$ we can find unions of almost disjoint rectangles A_ϵ and B_ϵ, such that $A_\epsilon \subset A$, $B_\epsilon \subset B$, with $m(A) \leq m(A_\epsilon) + \epsilon$ and $m(B) \leq m(B_\epsilon) + \epsilon$. Since $A + B \supset A_\epsilon + B_\epsilon$, the inequality (8) for A_ϵ and B_ϵ and a passage to a limit gives the desired result. From this, we can pass to the case where A and B are arbitrary compact sets, by noting first that $A + B$ is then compact, and that if we define $A^\epsilon = \{x : d(x, A) < \epsilon\}$, then A^ϵ are open, and $A^\epsilon \searrow A$ as $\epsilon \to 0$. With similar definitions for B^ϵ and $(A+B)^\epsilon$, we observe also that $A + B \subset A^\epsilon + B^\epsilon \subset (A+B)^{2\epsilon}$. Hence, letting $\epsilon \to 0$, we see that (8) for A^ϵ and B^ϵ implies the desired result for A and B. The general case, in which we assume that A, B, and $A + B$ are measurable, then follows by approximating A and B from inside by compact sets, as in (iii) of Theorem 3.4.

6 Exercises

1. Prove that the Cantor set $\mathcal{C}$ constructed in the text is totally disconnected and perfect. In other words, given two distinct points $x, y \in \mathcal{C}$, there is a point $z \notin \mathcal{C}$ that lies between x and y, and yet $\mathcal{C}$ has no isolated points.

[Hint: If $x, y \in \mathcal{C}$ and $|x - y| > 1/3^k$, then x and y belong to two different intervals in C_k. Also, given any $x \in \mathcal{C}$ there is an end-point y_k of some interval in C_k that satisfies $x \neq y_k$ and $|x - y_k| \leq 1/3^k$.]

2. The Cantor set $\mathcal{C}$ can also be described in terms of ternary expansions.

(a) Every number in $[0,1]$ has a ternary expansion

$$x = \sum_{k=1}^{\infty} a_k 3^{-k}, \qquad \text{where } a_k = 0, 1, \text{ or } 2.$$

Note that this decomposition is not unique since, for example, $1/3 = \sum_{k=2}^{\infty} 2/3^k$. Prove that $x \in \mathcal{C}$ if and only if x has a representation as above where every a_k is either 0 or 2.

(b) The **Cantor-Lebesgue function** is defined on $\mathcal{C}$ by

$$F(x) = \sum_{k=1}^{\infty} \frac{b_k}{2^k} \qquad \text{if } x = \textstyle\sum_{k=1}^{\infty} a_k 3^{-k}, \text{ where } b_k = a_k/2.$$

In this definition, we choose the expansion of x in which $a_k = 0$ or 2.

Show that F is well defined and continuous on $\mathcal{C}$, and moreover $F(0) = 0$ as well as $F(1) = 1$.

(c) Prove that $F : \mathcal{C} \to [0,1]$ is surjective, that is, for every $y \in [0,1]$ there exists $x \in \mathcal{C}$ such that $F(x) = y$.

(d) One can also extend F to be a continuous function on $[0,1]$ as follows. Note that if (a,b) is an open interval of the complement of $\mathcal{C}$, then $F(a) = F(b)$. Hence we may define F to have the constant value $F(a)$ in that interval.

A geometrical construction of F is described in Chapter 3.

3. Cantor sets of constant dissection. Consider the unit interval $[0,1]$, and let ξ be a fixed real number with $0 < \xi < 1$ (the case $\xi = 1/3$ corresponds to the Cantor set $\mathcal{C}$ in the text).

In stage 1 of the construction, remove the centrally situated open interval in $[0,1]$ of length ξ. In stage 2, remove two central intervals each of relative length ξ, one in each of the remaining intervals after stage 1, and so on.

Let $\mathcal{C}_\xi$ denote the set which remains after applying the above procedure indefinitely.[6]

(a) Prove that the complement of $\mathcal{C}_\xi$ in $[0,1]$ is the union of open intervals of total length equal to 1.

(b) Show directly that $m_*(\mathcal{C}_\xi) = 0$.

[Hint: After the kth stage, show that the remaining set has total length $= (1-\xi)^k$.]

4. Cantor-like sets. Construct a closed set $\hat{\mathcal{C}}$ so that at the kth stage of the construction one removes 2^{k-1} centrally situated open intervals each of length ℓ_k, with

$$\ell_1 + 2\ell_2 + \cdots + 2^{k-1}\ell_k < 1.$$

[6] The set we call $\mathcal{C}_\xi$ is sometimes denoted by $\mathcal{C}_{\frac{1-\xi}{2}}$.

(a) If ℓ_j are chosen small enough, then $\sum_{k=1}^{\infty} 2^{k-1}\ell_k < 1$. In this case, show that $m(\hat{\mathcal{C}}) > 0$, and in fact, $m(\hat{\mathcal{C}}) = 1 - \sum_{k=1}^{\infty} 2^{k-1}\ell_k$.

(b) Show that if $x \in \hat{\mathcal{C}}$, then there exists a sequence of points $\{x_n\}_{n=1}^{\infty}$ such that $x_n \notin \hat{\mathcal{C}}$, yet $x_n \to x$ and $x_n \in I_n$, where I_n is a sub-interval in the complement of $\hat{\mathcal{C}}$ with $|I_n| \to 0$.

(c) Prove as a consequence that $\hat{\mathcal{C}}$ is perfect, and contains no open interval.

(d) Show also that $\hat{\mathcal{C}}$ is uncountable.

5. Suppose E is a given set, and $\mathcal{O}_n$ is the open set:

$$\mathcal{O}_n = \{x :\ d(x, E) < 1/n\}.$$

Show:

(a) If E is compact, then $m(E) = \lim_{n\to\infty} m(\mathcal{O}_n)$.

(b) However, the conclusion in (a) may be false for E closed and unbounded; or E open and bounded.

6. Using translations and dilations, prove the following: Let B be a ball in $\mathbb{R}^d$ of radius r. Then $m(B) = v_d r^d$, where $v_d = m(B_1)$, and B_1 is the unit ball, $B_1 = \{x \in \mathbb{R}^d :\ |x| < 1\}$.

A calculation of the constant v_d is postponed until Exercise 14 in the next chapter.

7. If $\delta = (\delta_1, \ldots, \delta_d)$ is a d-tuple of positive numbers $\delta_i > 0$, and E is a subset of $\mathbb{R}^d$, we define δE by

$$\delta E = \{(\delta_1 x_1, \ldots, \delta_d x_d) :\ \text{where } (x_1, \ldots, x_d) \in E\}.$$

Prove that δE is measurable whenever E is measurable, and

$$m(\delta E) = \delta_1 \cdots \delta_d m(E).$$

8. Suppose L is a linear transformation of $\mathbb{R}^d$. Show that if E is a measurable subset of $\mathbb{R}^d$, then so is $L(E)$, by proceeding as follows:

(a) Note that if E is compact, so is $L(E)$. Hence if E is an F_σ set, so is $L(E)$.

(b) Because L automatically satisfies the inequality

$$|L(x) - L(x')| \leq M|x - x'|$$

for some M, we can see that L maps any cube of side length ℓ into a cube of side length $c_d M\ell$, with $c_d = 2\sqrt{d}$. Now if $m(E) = 0$, there is a collection of cubes $\{Q_j\}$ such that $E \subset \bigcup_j Q_j$, and $\sum_j m(Q_j) < \epsilon$. Thus $m_*(L(E)) \leq c'\epsilon$, and hence $m(L(E)) = 0$. Finally, use Corollary 3.5.

One can show that $m(L(E)) = |\det L|\, m(E)$; see Problem 4 in the next chapter.

9. Give an example of an open set $\mathcal{O}$ with the following property: the boundary of the closure of $\mathcal{O}$ has positive Lebesgue measure.

[Hint: Consider the set obtained by taking the union of open intervals which are deleted at the odd steps in the construction of a Cantor-like set.]

10. This exercise provides a construction of a decreasing sequence of positive continuous functions on the interval $[0, 1]$, whose pointwise limit is *not* Riemann integrable.

Let $\hat{\mathcal{C}}$ denote a Cantor-like set obtained from the construction detailed in Exercise 4, so that in particular $m(\hat{\mathcal{C}}) > 0$. Let F_1 denote a piecewise-linear and continuous function on $[0, 1]$, with $F_1 = 1$ in the complement of the first interval removed in the construction of $\hat{\mathcal{C}}$, $F_1 = 0$ at the center of this interval, and $0 \leq F_1(x) \leq 1$ for all x. Similarly, construct $F_2 = 1$ in the complement of the intervals in stage two of the construction of $\hat{\mathcal{C}}$, with $F_2 = 0$ at the center of these intervals, and $0 \leq F_2 \leq 1$. Continuing this way, let $f_n = F_1 \cdot F_2 \cdots F_n$ (see Figure 5).

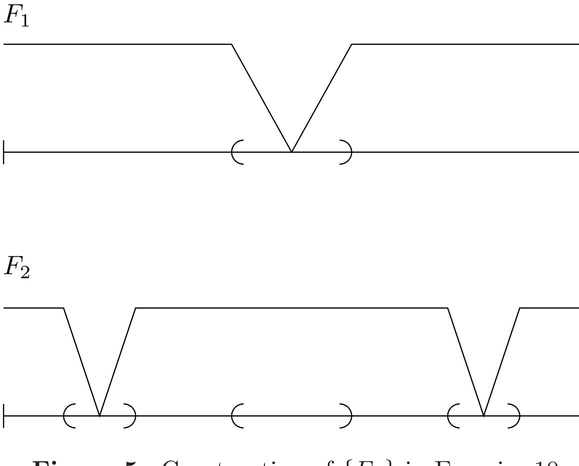


Figure 5. Construction of $\{F_n\}$ in Exercise 10

Prove the following:

(a) For all $n \geq 1$ and all $x \in [0, 1]$, one has $0 \leq f_n(x) \leq 1$ and $f_n(x) \geq f_{n+1}(x)$. Therefore, $f_n(x)$ converges to a limit as $n \to \infty$ which we denote by $f(x)$.

(b) The function f is discontinuous at every point of $\hat{\mathcal{C}}$.

[Hint: Note that $f(x) = 1$ if $x \in \hat{\mathcal{C}}$, and find a sequence of points $\{x_n\}$ so that $x_n \to x$ and $f(x_n) = 0$.]

Now $\int f_n(x)\, dx$ is decreasing, hence $\int f_n$ converges. However, a bounded function is Riemann integrable if and only if its set of discontinuities has measure zero.

(The proof of this fact, which is given in the Appendix of Book I, is outlined in Problem 4.) Since f is discontinuous on a set of positive measure, we find that f is not Riemann integrable.

11. Let A be the subset of $[0,1]$ which consists of all numbers which do not have the digit 4 appearing in their decimal expansion. Find $m(A)$.

12. Theorem 1.3 states that every open set in $\mathbb{R}$ is the disjoint union of open intervals. The analogue in $\mathbb{R}^d$, $d \geq 2$, is generally false. Prove the following:

(a) An open disc in $\mathbb{R}^2$ is not the disjoint union of open rectangles.

[Hint: What happens to the boundary of any of these rectangles?]

(b) An open connected set Ω is the disjoint union of open rectangles if and only if Ω is itself an open rectangle.

13. The following deals with G_δ and F_σ sets.

(a) Show that a closed set is a G_δ and an open set an F_σ.

[Hint: If F is closed, consider $\mathcal{O}_n = \{x : d(x,F) < 1/n\}$.]

(b) Give an example of an F_σ which is not a G_δ.

[Hint: This is more difficult; let F be a denumerable set that is dense.]

(c) Give an example of a Borel set which is not a G_δ nor an F_σ.

14. The purpose of this exercise is to show that covering by a *finite* number of intervals will not suffice in the definition of the outer measure m_*.

The **outer Jordan content** $J_*(E)$ of a set E in $\mathbb{R}$ is defined by

$$J_*(E) = \inf \sum_{j=1}^{N} |I_j|,$$

where the inf is taken over every *finite* covering $E \subset \bigcup_{j=1}^{N} I_j$, by intervals I_j.

(a) Prove that $J_*(E) = J_*(\overline{E})$ for every set E (here $\overline{E}$ denotes the closure of E).

(b) Exhibit a countable subset $E \subset [0,1]$ such that $J_*(E) = 1$ while $m_*(E) = 0$.

15. At the start of the theory, one might define the outer measure by taking coverings by rectangles instead of cubes. More precisely, we define

$$m_*^{\mathcal{R}}(E) = \inf \sum_{j=1}^{\infty} |R_j|,$$

where the inf is now taken over all countable coverings $E \subset \bigcup_{j=1}^{\infty} R_j$ by (closed) rectangles.

Show that this approach gives rise to the same theory of measure developed in the text, by proving that $m_*(E) = m_*^{\mathcal{R}}(E)$ for every subset E of $\mathbb{R}^d$.

[Hint: Use Lemma 1.1.]

16. The Borel-Cantelli lemma. Suppose $\{E_k\}_{k=1}^{\infty}$ is a countable family of measuable subsets of $\mathbb{R}^d$ and that

$$\sum_{k=1}^{\infty} m(E_k) < \infty.$$

Let

$$\begin{aligned} E &= \{x \in \mathbb{R}^d : x \in E_k, \text{ for infinitely many } k\} \\ &= \limsup_{k \to \infty}(E_k). \end{aligned}$$

(a) Show that E is measurable.

(b) Prove $m(E) = 0$.

[Hint: Write $E = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} E_k$.]

17. Let $\{f_n\}$ be a sequence of measurable functions on $[0,1]$ with $|f_n(x)| < \infty$ for a.e x. Show that there exists a sequence c_n of positive real numbers such that

$$\frac{f_n(x)}{c_n} \to 0 \qquad \text{a.e. } x$$

[Hint: Pick c_n such that $m(\{x : |f_n(x)/c_n| > 1/n\}) < 2^{-n}$, and apply the Borel-Cantelli lemma.]

18. Prove the following assertion: Every measurable function is the limit a.e. of a sequence of continuous functions.

19. Here are some observations regarding the set operation $A + B$.

(a) Show that if either A and B is open, then $A + B$ is open.

(b) Show that if A and B are closed, then $A + B$ is measurable.

(c) Show, however, that $A + B$ might not be closed even though A and B are closed.

[Hint: For (b) show that $A + B$ is an F_σ set.]

20. Show that there exist closed sets A and B with $m(A) = m(B) = 0$, but $m(A + B) > 0$:

(a) In $\mathbb{R}$, let $A = \mathcal{C}$ (the Cantor set), $B = \mathcal{C}/2$. Note that $A + B \supset [0,1]$.

(b) In $\mathbb{R}^2$, observe that if $A = I \times \{0\}$ and $B = \{0\} \times I$ (where $I = [0,1]$), then $A + B = I \times I$.

21. Prove that there is a continuous function that maps a Lebesgue measurable set to a non-measurable set.

[Hint: Consider a non-measurable subset of $[0,1]$, and its inverse image in $\mathcal{C}$ by the function F in Exercise 2.]

22. Let $\chi_{[0,1]}$ be the characteristic function of $[0,1]$. Show that there is no everywhere continuous function f on $\mathbb{R}$ such that

$$f(x) = \chi_{[0,1]}(x) \qquad \text{almost everywhere.}$$

23. Suppose $f(x,y)$ is a function on $\mathbb{R}^2$ that is separately continuous: for each fixed variable, f is continuous in the other variable. Prove that f is measurable on $\mathbb{R}^2$.

[Hint: Approximate f in the variable x by piecewise-linear functions f_n so that $f_n \to f$ pointwise.]

24. Does there exist an enumeration $\{r_n\}_{n=1}^\infty$ of the rationals, such that the complement of

$$\bigcup_{n=1}^\infty \left(r_n - \frac{1}{n}, r_n + \frac{1}{n} \right)$$

in $\mathbb{R}$ is non-empty?

[Hint: Find an enumeration where the only rationals outside of a fixed bounded interval take the form r_n, with $n = m^2$ for some integer m.]

25. An alternative definition of measurability is as follows: E is measurable if for every $\epsilon > 0$ there is a *closed* set F contained in E with $m_*(E - F) < \epsilon$. Show that this definition is equivalent with the one given in the text.

26. Suppose $A \subset E \subset B$, where A and B are measurable sets of finite measure. Prove that if $m(A) = m(B)$, then E is measurable.

27. Suppose E_1 and E_2 are a pair of compact sets in $\mathbb{R}^d$ with $E_1 \subset E_2$, and let $a = m(E_1)$ and $b = m(E_2)$. Prove that for any c with $a < c < b$, there is a compact set E with $E_1 \subset E \subset E_2$ and $m(E) = c$.

[Hint: As an example, if $d = 1$ and E is a measurable subset of $[0,1]$, consider $m(E \cap [0,t])$ as a function of t.]

28. Let E be a subset of $\mathbb{R}$ with $m_*(E) > 0$. Prove that for each $0 < \alpha < 1$, there exists an open interval I so that

$$m_*(E \cap I) \geq \alpha\, m_*(I).$$

Loosely speaking, this estimate shows that E contains almost a whole interval.

[Hint: Choose an open set $\mathcal{O}$ that contains E, and such that $m_*(E) \geq \alpha\, m_*(\mathcal{O})$. Write $\mathcal{O}$ as the countable union of disjoint open intervals, and show that one of these intervals must satisfy the desired property.]

29. Suppose E is a measurable subset of $\mathbb{R}$ with $m(E) > 0$. Prove that the **difference set** of E, which is defined by

$$\{z \in \mathbb{R} : \ z = x - y \text{ for some } x, y \in E\},$$

contains an open interval centered at the origin.

If E contains an interval, then the conclusion is straightforward. In general, one may rely on Exercise 28.

[Hint: Indeed, by Exercise 28, there exists an open interval I so that $m(E \cap I) \geq (9/10)\, m(I)$. If we denote $E \cap I$ by E_0, and suppose that the difference set of E_0 does not contain an open interval around the origin, then for arbitrarily small a the sets E_0, and $E_0 + a$ are disjoint. From the fact that $(E_0 \cup (E_0 + a)) \subset (I \cup (I + a))$ we get a contradiction, since the left-hand side has measure $2m(E_0)$, while the right-hand side has measure only slightly larger than $m(I)$.]

A more general formulation of this result is as follows.

30. If E and F are measurable, and $m(E) > 0$, $m(F) > 0$, prove that

$$E + F = \{x + y : x \in E,\ x \in F\}$$

contains an interval.

31. The result in Exercise 29 provides an alternate proof of the non-measurability of the set $\mathcal{N}$ studied in the text. In fact, we may also prove the non-measurability of a set in $\mathbb{R}$ that is very closely related to the set $\mathcal{N}$.

Given two real numbers x and y, we shall write as before that $x \sim y$ whenever the difference $x - y$ is rational. Let $\mathcal{N}^*$ denote a set that consists of one element in each equivalence class of $\sim$. Prove that $\mathcal{N}^*$ is non-measurable by using the result in Exercise 29.

[Hint: If $\mathcal{N}^*$ is measurable, then so are its translates $\mathcal{N}^*_n = \mathcal{N}^* + r_n$, where $\{r_n\}_{n=1}^{\infty}$ is an enumeration of $\mathbb{Q}$. How does this imply that $m(\mathcal{N}^*) > 0$? Can the difference set of $\mathcal{N}^*$ contain an open interval centered at the origin?]

32. Let $\mathcal{N}$ denote the non-measurable subset of $I = [0, 1]$ constructed at the end of Section 3.

(a) Prove that if E is a measurable subset of $\mathcal{N}$, then $m(E) = 0$.

(b) If G is a subset of $\mathbb{R}$ with $m_*(G) > 0$, prove that a subset of G is non-measurable.

[Hint: For (a) use the translates of E by the rationals.]

33. Let $\mathcal{N}$ denote the non-measurable set constructed in the text. Recall from the exercise above that measurable subsets of $\mathcal{N}$ have measure zero.

Show that the set $\mathcal{N}^c = I - \mathcal{N}$ satisfies $m_*(\mathcal{N}^c) = 1$, and conclude that if $E_1 = \mathcal{N}$ and $E_2 = \mathcal{N}^c$, then

$$m_*(E_1) + m_*(E_2) \neq m_*(E_1 \cup E_2),$$

although E_1 and E_2 are disjoint.

[Hint: To prove that $m_*(\mathcal{N}^c) = 1$, argue by contradiction and pick a measurable set U such that $U \subset I$, $\mathcal{N}^c \subset U$ and $m_*(U) < 1 - \epsilon$.]

34. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be any two Cantor sets (constructed in Exercise 3). Show that there exists a function $F : [0,1] \to [0,1]$ with the following properties:

(i) F is continuous and bijective,

(ii) F is monotonically increasing,

(iii) F maps $\mathcal{C}_1$ surjectively onto $\mathcal{C}_2$.

[Hint: Copy the construction of the standard Cantor-Lebesgue function.]

35. Give an example of a measurable function f and a continuous function Φ so that $f \circ \Phi$ is non-measurable.

[Hint: Let $\Phi : \mathcal{C}_1 \to \mathcal{C}_2$ as in Exercise 34, with $m(\mathcal{C}_1) > 0$ and $m(\mathcal{C}_2) = 0$. Let $N \subset \mathcal{C}_1$ be non-measurable, and take $f = \chi_{\Phi(N)}$.]

Use the construction in the hint to show that there exists a Lebesgue measurable set that is not a Borel set.

36. This exercise provides an example of a measurable function f on $[0,1]$ such that every function g equivalent to f (in the sense that f and g differ only on a set of measure zero) is discontinuous at *every* point.

(a) Construct a measurable set $E \subset [0,1]$ such that for any non-empty open sub-interval I in $[0,1]$, both sets $E \cap I$ and $E^c \cap I$ have positive measure.

(b) Show that $f = \chi_E$ has the property that whenever $g(x) = f(x)$ a.e x, then g must be discontinuous at every point in $[0,1]$.

[Hint: For the first part, consider a Cantor-like set of positive measure, and add in each of the intervals that are omitted in the first step of its construction, another Cantor-like set. Continue this procedure indefinitely.]

37. Suppose Γ is a curve $y = f(x)$ in $\mathbb{R}^2$, where f is continuous. Show that $m(\Gamma) = 0$.

[Hint: Cover Γ by rectangles, using the uniform continuity of f.]

38. Prove that $(a+b)^\gamma \geq a^\gamma + b^\gamma$ whenever $\gamma \geq 1$ and $a, b \geq 0$. Also, show that the reverse inequality holds when $0 \leq \gamma \leq 1$.
[Hint: Integrate the inequality between $(a+t)^{\gamma-1}$ and $t^{\gamma-1}$ from 0 to b.]

39. Establish the inequality

$$\frac{x_1 + \cdots + x_d}{d} \geq (x_1 \cdots x_d)^{1/d} \qquad \text{for all } x_j \geq 0,\ j = 1, \ldots, d \tag{10}$$

by using backward induction as follows:

(a) The inequality is true whenever d is a power of 2 ($d = 2^k$, $k \geq 1$).

(b) If (10) holds for some integer $d \geq 2$, then it must hold for $d-1$, that is, one has $(y_1 + \cdots + y_{d-1})/(d-1) \geq (y_1 \cdots y_{d-1})^{1/(d-1)}$ for all $y_j \geq 0$, with $j = 1, \ldots, d-1$.

[Hint: For (a), if $k \geq 2$, write $(x_1 + \cdots + x_{2^k})/2^k$ as $(A+B)/2$, where $A = (x_1 + \cdots + x_{2^{k-1}})/2^{k-1}$, and apply the inequality when $d = 2$. For (b), apply the inequality to $x_1 = y_1, \ldots, x_{d-1} = y_{d-1}$ and $x_d = (y_1 + \cdots + y_{d-1})/(d-1)$.]

7 Problems

1. Given an irrational x, one can show (using the pigeon-hole principle, for example) that there exists infinitely many fractions p/q, with relatively prime integers p and q such that

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{q^2}.$$

However, prove that the set of those $x \in \mathbb{R}$ such that there exist infinitely many fractions p/q, with relatively prime integers p and q such that

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{q^3} \qquad (\text{or } \leq 1/q^{2+\epsilon}),$$

is a set of measure zero.
[Hint: Use the Borel-Cantelli lemma.]

2. Any open set Ω can be written as the union of closed cubes, so that $\Omega = \bigcup Q_j$ with the following properties

(i) The Q_j's have disjoint interiors.

(ii) $d(Q_j, \Omega^c) \approx$ side length of Q_j. This means that there are positive constants c and C so that $c \leq d(Q_j, \Omega^c)/\ell(Q_j) \leq C$, where $\ell(Q_j)$ denotes the side length of Q_j.

3. Find an example of a measurable subset C of $[0,1]$ such that $m(C) = 0$, yet the difference set of C contains a non-trivial interval centered at the origin. Compare with the result in Exercise 29.

[Hint: Pick the Cantor set $C = \mathcal{C}$. For a fixed $a \in [-1,1]$, consider the line $y = x + a$ in the plane, and copy the construction of the Cantor set, but in the cube $Q = [0,1] \times [0,1]$. First, remove all but four closed cubes of side length 1/3, one at each corner of Q; then, repeat this procedure in each of the remaining cubes (see Figure 6). The resulting set is sometimes called a Cantor dust. Use the property of nested compact sets to show that the line intersects this Cantor dust.]

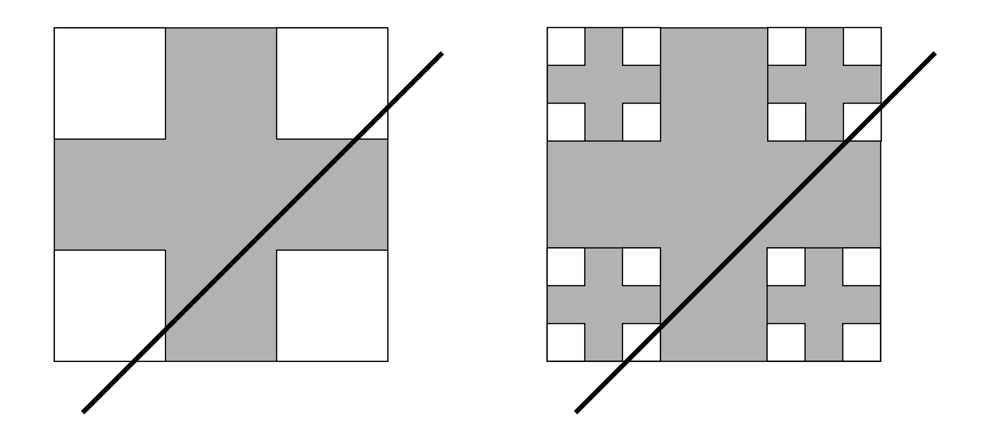


Figure 6. Construction of the Cantor dust

4. Complete the following outline to prove that a bounded function on an interval $[a,b]$ is Riemann integrable if and only if its set of discontinuities has measure zero. This argument is given in detail in the appendix to Book I.

Let f be a *bounded* function on a compact interval J, and let $I(c,r)$ denote the open interval centered at c of radius $r > 0$. Let $\text{osc}(f,c,r) = \sup |f(x) - f(y)|$, where the supremum is taken over all $x, y \in J \cap I(c,r)$, and define the oscillation of f at c by $\text{osc}(f,c) = \lim_{r \to 0} \text{osc}(f,c,r)$. Clearly, f is continuous at $c \in J$ if and only if $\text{osc}(f,c) = 0$.

Prove the following assertions:

(a) For every $\epsilon > 0$, the set of points c in J such that $\text{osc}(f,c) \geq \epsilon$ is compact.

(b) If the set of discontinuities of f has measure 0, then f is Riemann integrable.

[Hint: Given $\epsilon > 0$ let $A_\epsilon = \{c \in J : \text{osc}(f,c) \geq \epsilon\}$. Cover A_ϵ by a finite number of open intervals whose total length is $\leq \epsilon$. Select an appropriate partition of J and estimate the difference between the upper and lower sums of f over this partition.]

(c) Conversely, if f is Riemann integrable on J, then its set of discontinuities has measure 0.

[Hint: The set of discontinuities of f is contained in $\bigcup_n A_{1/n}$. Choose a partition P such that $U(f, P) - L(f, P) < \epsilon/n$. Show that the total length of the intervals in P whose interior intersect $A_{1/n}$ is $\leq \epsilon$.]

5. Suppose E is measurable with $m(E) < \infty$, and

$$E = E_1 \cup E_2, \qquad E_1 \cap E_2 = \emptyset.$$

If $m(E) = m_*(E_1) + m_*(E_2)$, then E_1 and E_2 are measurable.

In particular, if $E \subset Q$, where Q is a finite cube, then E is measurable if and only if $m(Q) = m_*(E) + m_*(Q - E)$.

6.* The fact that the axiom of choice and the well-ordering principle are equivalent is a consequence of the following considerations.

One begins by defining a partial ordering on a set E to be a binary relation $\leq$ on the set E that satisfies:

(i) $x \leq x$ for all $x \in E$.

(ii) If $x \leq y$ and $y \leq x$, then $x = y$.

(iii) If $x \leq y$ and $y \leq z$, then $x \leq z$.

If in addition $x \leq y$ or $y \leq x$ whenever $x, y \in E$, then $\leq$ is a linear ordering of E.

The axiom of choice and the well-ordering principle are then logically equivalent to the Hausdorff maximal principle:

Every non-empty partially ordered set has a (non-empty) maximal linearly ordered subset.

In other words, if E is partially ordered by $\leq$, then E contains a non-empty subset F which is linearly ordered by $\leq$ and such that if F is contained in a set G also linearly ordered by $\leq$, then $F = G$.

An application of the Hausdorff maximal principle to the collection of all well-orderings of subsets of E implies the well-ordering principle for E. However, the proof that the axiom of choice implies the Hausdorff maximal principle is more complicated.

7.* Consider the curve $\Gamma = \{y = f(x)\}$ in $\mathbb{R}^2$, $0 \leq x \leq 1$. Assume that f is twice continuously differentiable in $0 \leq x \leq 1$. Then show that $m(\Gamma + \Gamma) > 0$ if and only if $\Gamma + \Gamma$ contains an open set, if and only if f is not linear.

8.* Suppose A and B are open sets of finite positive measure. Then we have equality in the Brunn-Minkowski inequality (8) if and only if A and B are convex and similar, that is, there are a $\delta > 0$ and an $h \in \mathbb{R}^d$ such that

$$A = \delta B + h.$$

2 Integration Theory

> ...amongst the many definitions that have been successively proposed for the integral of real-valued functions of a real variable, I have retained only those which, in my opinion, are indispensable to understand the transformations undergone by the problem of integration, and to capture the relationship between the notion of area, so simple in appearance, and certain more complicated analytical definitions of the integral.
>
> One might ask if there is sufficient interest to occupy oneself with such complications, and if it is not better to restrict oneself to the study of functions that necessitate only simple definitions.... As we shall see in this course, we would then have to renounce the possibility of resolving many problems posed long ago, and which have simple statements. It is to solve these problems, and not for love of complications, that I have introduced in this book a definition of the integral more general than that of Riemann.
>
> *H. Lebesgue,* 1903

1 The Lebesgue integral: basic properties and convergence theorems

The general notion of the Lebesgue integral on $\mathbb{R}^d$ will be defined in a step-by-step fashion, proceeding successively to increasingly larger families of functions. At each stage we shall see that the integral satisfies elementary properties such as linearity and monotonicity, and we prove appropriate convergence theorems that amount to interchanging the integral with limits. At the end of the process we shall have achieved a general theory of integration that will be decisive in the study of further problems.

We proceed in four stages, by progressively integrating:

1. Simple functions
2. Bounded functions supported on a set of finite measure
3. Non-negative functions

4. Integrable functions (the general case).

We emphasize from the onset that all functions are assumed to be measurable. At the beginning we also consider only finite-valued functions which take on real values. Later we shall also consider extended-valued functions, and also complex-valued functions.

Stage one: simple functions

Recall from the previous chapter that a simple function φ is a finite sum

$$\varphi(x) = \sum_{k=1}^{N} a_k \chi_{E_k}(x), \tag{1}$$

where the E_k are measurable sets of finite measure and the a_k are constants. A complication that arises from this definition is that a simple function can be written in a multitude of ways as such finite linear combinations; for example, $0 = \chi_E - \chi_E$ for any measurable set E of finite measure. Fortunately, there is an unambiguous choice for the representation of a simple function, which is natural and useful in applications.

The **canonical form** of φ is the unique decomposition as in (1), where the numbers a_k are distinct and non-zero, and the sets E_k are disjoint.

Finding the canonical form of φ is straightforward: since φ can take only finitely many distinct and non-zero values, say $c_1, \ldots, c_M$, we may set $F_k = \{x : \varphi(x) = c_k\}$, and note that the sets F_k are disjoint. Therefore $\varphi = \sum_{k=1}^{M} c_k \chi_{F_k}$ is the desired canonical form of φ.

If φ is a simple function with canonical form $\varphi(x) = \sum_{k=1}^{M} c_k \chi_{F_k}(x)$, then we *define* the **Lebesgue integral** of φ by

$$\int_{\mathbb{R}^d} \varphi(x)\, dx = \sum_{k=1}^{M} c_k m(F_k).$$

If E is a measurable subset of $\mathbb{R}^d$ with finite measure, then $\varphi(x)\chi_E(x)$ is also a simple function, and we *define*

$$\int_E \varphi(x)\, dx = \int \varphi(x)\chi_E(x)\, dx.$$

To emphasize the choice of the Lebesgue measure m in the definition of the integral, one sometimes writes

$$\int_{\mathbb{R}^d} \varphi(x)\, dm(x)$$

for the Lebesgue integral of φ. In fact, as a matter of convenience, we shall often write $\int \varphi(x)\,dx$ or simply $\int \varphi$ for the integral of φ over $\mathbb{R}^d$.

Proposition 1.1 *The integral of simple functions defined above satisfies the following properties:*

(i) *Independence of the representation. If $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$ is any representation of φ, then*

$$\int \varphi = \sum_{k=1}^{N} a_k m(E_k).$$

(ii) *Linearity. If φ and ψ are simple, and $a, b \in \mathbb{R}$, then*

$$\int (a\varphi + b\psi) = a \int \varphi + b \int \psi.$$

(iii) *Additivity. If E and F are disjoint subsets of $\mathbb{R}^d$ with finite measure, then*

$$\int_{E \cup F} \varphi = \int_E \varphi + \int_F \varphi.$$

(iv) *Monotonicity. If $\varphi \leq \psi$ are simple, then*

$$\int \varphi \leq \int \psi.$$

(v) *Triangle inequality. If φ is a simple function, then so is $|\varphi|$, and*

$$\left| \int \varphi \right| \leq \int |\varphi|.$$

Proof. The only conclusion that is a little tricky is the first, which asserts that the integral of a simple function can be calculated by using any of its decompositions as a linear combination of characteristic functions.

Suppose that $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$, where we assume that the sets E_k are disjoint, but we do not suppose that the numbers a_k are distinct and non-zero. For each distinct non-zero value a among the $\{a_k\}$ we define $E'_a = \bigcup E_k$, where the union is taken over those indices k such that $a_k = a$. Note then that the sets E'_a are disjoint, and $m(E'_a) = \sum m(E_k)$, where

the sum is taken over the same set of k's. Then clearly $\varphi = \sum a\chi_{E'_a}$, where the sum is over the distinct non-zero values of $\{a_k\}$. Thus

$$\int \varphi = \sum am(E'_a) = \sum_{k=1}^{N} a_k m(E_k).$$

Next, suppose $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$, where we no longer assume that the E_k are disjoint. Then we can "refine" the decomposition $\bigcup_{k=1}^{N} E_k$ by finding sets $E_1^*,\ E_2^*, \ldots, E_n^*$ with the property that $\bigcup_{k=1}^{N} E_k = \bigcup_{j=1}^{n} E_j^*$; the sets E_j^* $(j = 1, \ldots, n)$ are mutually disjoint; and for each k, $E_k = \bigcup E_j^*$, where the union is taken over those E_j^* that are contained in E_k. (A proof of this elementary fact can be found in Exercise 1.) For each j, let now $a_j^* = \sum a_k$, with the summation taken over all k such that E_k contains E_j^*. Then clearly $\varphi = \sum_{j=1}^{n} a_j^* \chi_{E_j^*}$. However, this is a decomposition already dealt with above because the E_j^* are disjoint. Thus

$$\int \varphi = \sum a_j^* m(E_j^*) = \sum \sum_{E_k \supset E_j^*} a_k m(E_j^*) = \sum a_k m(E_k),$$

and conclusion (i) is established.

Conclusion (ii) follows by using any representation of φ and ψ, and the obvious linearity of (i).

For the additivity over sets, one must note that if E and F are disjoint, then

$$\chi_{E \cup F} = \chi_E + \chi_F,$$

and we may use the linearity of the integral to see that $\int_{E \cup F} \varphi = \int_E \varphi + \int_F \varphi$.

If $\eta \geq 0$ is a simple function, then its canonical form is everywhere non-negative, and therefore $\int \eta \geq 0$ by the definition of the integral. Applying this argument to $\psi - \varphi$ gives the desired monotonicity property.

Finally, for the triangle inequality, it suffices to write φ in its canonical form $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$ and observe that

$$|\varphi| = \sum_{k=1}^{N} |a_k| \chi_{E_k}(x).$$

Therefore, by the triangle inequality applied to the definition of the integral, one sees that

$$\left| \int \varphi \right| = \left| \sum_{k=1}^{N} a_k m(E_k) \right| \leq \sum_{k=1}^{N} |a_k| m(E_k) = \int |\varphi|.$$

Incidentally, it is worthwhile to point out the following easy fact: whenever f and g are a pair of simple functions that agree almost everywhere, then $\int f = \int g$. The identity of the integrals of two functions that agree almost everywhere will continue to hold for the successive definitions of the integral that follow.

Stage two: bounded functions supported on a set of finite measure

The **support** of a measurable function f is defined to be the set of all points where f does not vanish,

$$\text{supp}(f) = \{x : f(x) \neq 0\}.$$

We shall also say that f is **supported** on a set E, if $f(x) = 0$ whenever $x \notin E$.

Since f is measurable, so is the set $\text{supp}(f)$. We shall next be interested in those bounded measurable functions that have $m(\text{supp}(f)) < \infty$.

An important result in the previous chapter (Theorem 4.2) states the following: if f is a function bounded by M and supported on a set E, then there exists a sequence $\{\varphi_n\}$ of simple functions, with each φ_n bounded by M and supported on E, and such that

$$\varphi_n(x) \to f(x) \quad \text{for all } x.$$

The key lemma that follows allows us to define the integral for the class of bounded functions supported on sets of finite measure.

Lemma 1.2 *Let f be a bounded function supported on a set E of finite measure. If $\{\varphi_n\}_{n=1}^{\infty}$ is any sequence of simple functions bounded by M, supported on E, and with $\varphi_n(x) \to f(x)$ for a.e. x, then:*

(i) *The limit* $\displaystyle\lim_{n\to\infty} \int \varphi_n$ *exists.*

(ii) *If $f = 0$ a.e., then the limit* $\displaystyle\lim_{n\to\infty} \int \varphi_n$ *equals* 0.

Proof. The assertions of the lemma would be nearly obvious if we had that φ_n converges to f uniformly on E. Instead, we recall one of Littlewood's principles, which states that the convergence of a sequence of measurable functions is "nearly" uniform. The precise statement lying behind this principle is Egorov's theorem, which we proved in Chapter 1, and which we apply here.

Since the measure of E is finite, given $\epsilon > 0$ Egorov's theorem guarantees the existence of a (closed) measurable subset A_ϵ of E such that $m(E - A_\epsilon) \leq \epsilon$, and $\varphi_n \to f$ uniformly on A_ϵ. Therefore, setting $I_n = \int \varphi_n$ we have that

$$\begin{aligned} |I_n - I_m| &\leq \int_E |\varphi_n(x) - \varphi_m(x)|\, dx \\ &= \int_{A_\epsilon} |\varphi_n(x) - \varphi_m(x)|\, dx + \int_{E - A_\epsilon} |\varphi_n(x) - \varphi_m(x)|\, dx \\ &\leq \int_{A_\epsilon} |\varphi_n(x) - \varphi_m(x)|\, dx + 2M\, m(E - A_\epsilon) \\ &\leq \int_{A_\epsilon} |\varphi_n(x) - \varphi_m(x)|\, dx + 2M\epsilon. \end{aligned}$$

By the uniform convergence, one has, for all $x \in A_\epsilon$ and all large n and m, the estimate $|\varphi_n(x) - \varphi_m(x)| < \epsilon$, so we deduce that

$$|I_n - I_m| \leq m(E)\epsilon + 2M\epsilon \quad \text{for all large } n \text{ and } m.$$

Since ϵ is arbitrary and $m(E) < \infty$, this proves that $\{I_n\}$ is a Cauchy sequence and hence converges, as desired.

For the second part, we note that if $f = 0$, we may repeat the argument above to find that $|I_n| \leq m(E)\epsilon + M\epsilon$, which yields $\lim_{n\to\infty} I_n = 0$, as was to be shown.

Using Lemma 1.2 we can now turn to the integration of bounded functions that are supported on sets of finite measure. For such a function f we *define* its **Lebesgue integral** by

$$\int f(x)\, dx = \lim_{n\to\infty} \int \varphi_n(x)\, dx,$$

where $\{\varphi_n\}$ is any sequence of simple functions satisfying: $|\varphi_n| \leq M$, each φ_n is supported on the support of f, and $\varphi_n(x) \to f(x)$ for a.e. x as n tends to infinity. By the previous lemma, we know that this limit exists.

Next, we must first show that $\int f$ is independent of the limiting sequence $\{\varphi_n\}$ used, in order for the integral to be well-defined. Therefore, suppose that $\{\psi_n\}$ is another sequence of simple functions that is bounded by M, supported on $\text{supp}(f)$, and such that $\psi_n(x) \to f(x)$ for a.e. x as n tends to infinity. Then, if $\eta_n = \varphi_n - \psi_n$, the sequence $\{\eta_n\}$ consists of simple functions bounded by $2M$, supported on a set of finite measure, and such that $\eta_n \to 0$ a.e. as n tends to infinity. We may

therefore conclude, by the second part of the lemma, that $\int \eta_n \to 0$ as n tends to infinity. Consequently, the two limits

$$\lim_{n\to\infty} \int \varphi_n(x)\,dx \quad \text{and} \quad \lim_{n\to\infty} \int \psi_n(x)\,dx$$

(which exist by the lemma) are indeed equal.

If E is a subset of $\mathbb{R}^d$ with finite measure, and f is bounded with $m(\mathrm{supp}(f)) < \infty$, then it is natural to define

$$\int_E f(x)\,dx = \int f(x)\chi_E(x)\,dx.$$

Clearly, if f is itself simple, then $\int f$ as defined above coincides with the integral of simple functions studied earlier. This extension of the definition of integration also satisfies all the basic properties of the integral of simple functions.

Proposition 1.3 *Suppose f and g are bounded functions supported on sets of finite measure. Then the following properties hold.*

(i) *Linearity. If $a, b \in \mathbb{R}$, then*

$$\int (af + bg) = a\int f + b\int g.$$

(ii) *Additivity. If E and F are disjoint subsets of $\mathbb{R}^d$, then*

$$\int_{E\cup F} f = \int_E f + \int_F f.$$

(iii) *Monotonicity. If $f \le g$, then*

$$\int f \le \int g.$$

(iv) *Triangle inequality. $|f|$ is also bounded, supported on a set of finite measure, and*

$$\left|\int f\right| \le \int |f|.$$

All these properties follow by using approximations by simple functions, and the properties of the integral of simple functions given in Proposition 1.1.

We are now in a position to prove the first important convergence theorem.

Theorem 1.4 (Bounded convergence theorem) *Suppose that* $\{f_n\}$ *is a sequence of measurable functions that are all bounded by* M*, are supported on a set* E *of finite measure, and* $f_n(x) \to f(x)$ *a.e.* x *as* $n \to \infty$*. Then* f *is measurable, bounded, supported on* E *for a.e.* x*, and*

$$\int |f_n - f| \to 0 \qquad \textit{as } n \to \infty.$$

Consequently,

$$\int f_n \to \int f \qquad \textit{as } n \to \infty.$$

Proof. From the assumptions one sees at once that f is bounded by M almost everywhere and vanishes outside E, except possibly on a set of measure zero. Clearly, the triangle inequality for the integral implies that it suffices to prove that $\int |f_n - f| \to 0$ as n tends to infinity.

The proof is a reprise of the argument in Lemma 1.2. Given $\epsilon > 0$, we may find, by Egorov's theorem, a measurable subset A_ϵ of E such that $m(E - A_\epsilon) \leq \epsilon$ and $f_n \to f$ uniformly on A_ϵ. Then, we know that for all sufficiently large n we have $|f_n(x) - f(x)| \leq \epsilon$ for all $x \in A_\epsilon$. Putting these facts together yields

$$\begin{aligned}\int |f_n(x) - f(x)|\,dx &\leq \int_{A_\epsilon} |f_n(x) - f(x)|\,dx + \int_{E-A_\epsilon} |f_n(x) - f(x)|\,dx \\ &\leq \epsilon m(E) + 2M\,m(E - A_\epsilon)\end{aligned}$$

for all large n. Since ϵ is arbitrary, the proof of the theorem is complete.

We note that the above convergence theorem is a statement about the interchange of an integral and a limit, since its conclusion simply says

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n.$$

A useful observation that we can make at this point is the following: if $f \geq 0$ is bounded and supported on a set of finite measure E and $\int f = 0$,

then $f = 0$ almost everywhere. Indeed, if for each integer $k \geq 1$ we set $E_k = \{x \in E : f(x) \geq 1/k\}$, then the fact that $k^{-1}\chi_{E_k}(x) \leq f(x)$ implies

$$k^{-1}m(E_k) \leq \int f,$$

by monotonicity of the integral. Thus $m(E_k) = 0$ for all k, and since $\{x : f(x) > 0\} = \bigcup_{k=1}^{\infty} E_k$, we see that $f = 0$ almost everywhere.

Return to Riemann integrable functions

We shall now show that Riemann integrable functions are also Lebesgue integrable. When we combine this with the bounded convergence theorem we have just proved, we see that Lebesgue integration resolves the second problem in the Introduction.

Theorem 1.5 *Suppose f is Riemann integrable on the closed interval $[a,b]$. Then f is measurable, and*

$$\int_{[a,b]}^{\mathcal{R}} f(x)\,dx = \int_{[a,b]}^{\mathcal{L}} f(x)\,dx,$$

where the integral on the left-hand side is the standard Riemann integral, and that on the right-hand side is the Lebesgue integral.

Proof. By definition, a Riemann integrable function is bounded, say $|f(x)| \leq M$, so we need to prove that f is measurable, and then establish the equality of integrals.

Again, by definition of Riemann integrability,[1] we may construct two sequences of step functions $\{\varphi_k\}$ and $\{\psi_k\}$ that satisfy the following properties: $|\varphi_k(x)| \leq M$ and $|\psi_k(x)| \leq M$ for all $x \in [a,b]$ and $k \geq 1$,

$$\varphi_1(x) \leq \varphi_2(x) \leq \cdots \leq f \leq \cdots \leq \psi_2(x) \leq \psi_1(x),$$

and

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k(x)\,dx = \lim_{k\to\infty} \int_{[a,b]}^{\mathcal{R}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{R}} f(x)\,dx. \tag{2}$$

Several observations are in order. First, it follows immediately from their definition that for step functions the Riemann and Lebesgue integrals agree; therefore

(3)

$$\int_{[a,b]}^{\mathcal{R}} \varphi_k(x)\,dx = \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)\,dx \quad \text{and} \quad \int_{[a,b]}^{\mathcal{R}} \psi_k(x)\,dx = \int_{[a,b]}^{\mathcal{L}} \psi_k(x)\,dx$$

[1] See also Section 1 of the Appendix in Book I.

for all $k \geq 1$. Next, if we let

$$\tilde{\varphi}(x) = \lim_{k\to\infty} \varphi_k(x) \quad \text{and} \quad \tilde{\psi}(x) = \lim_{k\to\infty} \psi_k(x),$$

we have $\tilde{\varphi} \leq f \leq \tilde{\psi}$. Moreover, both $\tilde{\varphi}$ and $\tilde{\psi}$ are measurable (being the limit of step functions), and the bounded convergence theorem yields

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)\, dx = \int_{[a,b]}^{\mathcal{L}} \tilde{\varphi}(x)\, dx$$

and

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \psi_k(x)\, dx = \int_{[a,b]}^{\mathcal{L}} \tilde{\psi}(x)\, dx.$$

This together with (2) and (3) yields

$$\int_{[a,b]}^{\mathcal{L}} (\tilde{\psi}(x) - \tilde{\varphi}(x))\, dx = 0,$$

and since $\psi_k - \varphi_k \geq 0$, we must have $\tilde{\psi} - \tilde{\varphi} \geq 0$. By the observation following the proof of the bounded convergence theorem, we conclude that $\tilde{\psi} - \tilde{\varphi} = 0$ a.e., and therefore $\tilde{\varphi} = \tilde{\psi} = f$ a.e., which proves that f is measurable. Finally, since $\varphi_k \to f$ almost everywhere, we have (by definition)

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)\, dx = \int_{[a,b]}^{\mathcal{L}} f(x)\, dx,$$

and by (2) and (3) we see that $\int_{[a,b]}^{\mathcal{R}} f(x)\, dx = \int_{[a,b]}^{\mathcal{L}} f(x)\, dx$, as desired.

Stage three: non-negative functions

We proceed with the integrals of functions that are measurable and non-negative but not necessarily bounded. It will be important to allow these functions to be extended-valued, that is, these functions may take on the value $+\infty$ (on a measurable set). We recall in this connection the convention that one defines the supremum of a set of positive numbers to be $+\infty$ if the set is unbounded.

In the case of such a function f we *define* its (extended) **Lebesgue integral** by

$$\int f(x)\, dx = \sup_g \int g(x)\, dx,$$

where the supremum is taken over all measurable functions g such that $0 \leq g \leq f$, and where g is bounded and supported on a set of finite measure.

With the above definition of the integral, there are only two possible cases; the supremum is either finite, or infinite. In the first case, when $\int f(x)\,dx < \infty$, we shall say that f is **Lebesgue integrable** or simply **integrable**.

Clearly, if E is any measurable subset of $\mathbb{R}^d$, and $f \geq 0$, then $f\chi_E$ is also positive, and we define

$$\int_E f(x)\,dx = \int f(x)\chi_E(x)\,dx.$$

Simple examples of functions on $\mathbb{R}^d$ that are integrable (or non-integrable) are given by

$$f_a(x) = \begin{cases} |x|^{-a} & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| > 1. \end{cases}$$

$$F_a(x) = \frac{1}{1+|x|^a}, \qquad \text{all } x \in \mathbb{R}^d.$$

Then f_a is integrable exactly when $a < d$, while F_a is integrable exactly when $a > d$. See the discussion following Corollary 1.10 and also Exercise 10.

Proposition 1.6 *The integral of non-negative measurable functions enjoys the following properties:*

(i) *Linearity. If $f, g \geq 0$, and a, b are positive real numbers, then*

$$\int (af + bg) = a\int f + b\int g.$$

(ii) *Additivity. If E and F are disjoint subsets of $\mathbb{R}^d$, and $f \geq 0$, then*

$$\int_{E\cup F} f = \int_E f + \int_F f.$$

(iii) *Monotonicity. If $0 \leq f \leq g$, then*

$$\int f \leq \int g.$$

(iv) *If g is integrable and $0 \leq f \leq g$, then f is integrable.*

(v) *If f is integrable, then $f(x) < \infty$ for almost every x.*

(vi) *If $\int f = 0$, then $f(x) = 0$ for almost every x.*

Proof. Of the first four assertions, only (i) is not an immediate consequence of the definitions, and to prove it we argue as follows. We take $a = b = 1$ and note that if $\varphi \leq f$ and $\psi \leq g$, where both φ and ψ are bounded and supported on sets of finite measure, then $\varphi + \psi \leq f + g$, and $\varphi + \psi$ is also bounded and supported on a set of finite measure. Consequently

$$\int f + \int g \leq \int (f + g).$$

To prove the reverse inequality, suppose η is bounded and supported on a set of finite measure, and $\eta \leq f + g$. If we define $\eta_1(x) = \min(f(x), \eta(x))$ and $\eta_2 = \eta - \eta_1$, we note that

$$\eta_1 \leq f \quad \text{and} \quad \eta_2 \leq g.$$

Moreover both η_1, η_2 are bounded and supported on sets of finite measure. Hence

$$\int \eta = \int (\eta_1 + \eta_2) = \int \eta_1 + \int \eta_2 \leq \int f + \int g.$$

Taking the supremum over η yields the required inequality.

To prove the conclusion (v) we argue as follows. Suppose $E_k = \{x : f(x) \geq k\}$, and $E_\infty = \{x : f(x) = \infty\}$. Then

$$\int f \geq \int \chi_{E_k} f \geq k m(E_k),$$

hence $m(E_k) \to 0$ as $k \to \infty$. Since $E_k \searrow E_\infty$, Corollary 3.3 in the previous chapter implies that $m(E_\infty) = 0$.

The proof of (vi) is the same as the observation following Theorem 1.4.

We now turn our attention to some important convergence theorems for the class of non-negative measurable functions. To motivate the results that follow, we ask the following question: Suppose $f_n \geq 0$ and $f_n(x) \to f(x)$ for almost every x. Is it true that $\int f_n \, dx \to \int f \, dx$? Unfortunately, the example that follows provides a negative answer to this,

and shows that we must change our formulation of the question to obtain a positive convergence result.

Let

$$f_n(x) = \begin{cases} n & \text{if } 0 < x < 1/n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f_n(x) \to 0$ for all x, yet $\int f_n(x)\,dx = 1$ for all n. In this particular example, the limit of the integrals is greater than the integral of the limit function. This turns out to be the case in general, as we shall see now.

Lemma 1.7 (Fatou) *Suppose $\{f_n\}$ is a sequence of measurable functions with $f_n \geq 0$. If $\lim_{n\to\infty} f_n(x) = f(x)$ for a.e. x, then*

$$\int f \leq \liminf_{n\to\infty} \int f_n.$$

Proof. Suppose $0 \leq g \leq f$, where g is bounded and supported on a set E of finite measure. If we set $g_n(x) = \min(g(x), f_n(x))$, then g_n is measurable, supported on E, and $g_n(x) \to g(x)$ a.e., so by the bounded convergence theorem

$$\int g_n \to \int g.$$

By construction, we also have $g_n \leq f_n$, so that $\int g_n \leq \int f_n$, and therefore

$$\int g \leq \liminf_{n\to\infty} \int f_n.$$

Taking the supremum over all g yields the desired inequality.

In particular, we do not exclude the cases $\int f = \infty$, or $\liminf_{n\to\infty} f_n = \infty$.

We can now immediately deduce the following series of corollaries.

Corollary 1.8 *Suppose f is a non-negative measurable function, and $\{f_n\}$ a sequence of non-negative measurable functions with $f_n(x) \leq f(x)$ and $f_n(x) \to f(x)$ for almost every x. Then*

$$\lim_{n\to\infty} \int f_n = \int f.$$

Proof. Since $f_n(x) \leq f(x)$ a.e x, we necessarily have $\int f_n \leq \int f$ for all n; hence

$$\limsup_{n\to\infty} \int f_n \leq \int f.$$

This inequality combined with Fatou's lemma proves the desired limit.

In particular, we can now obtain a basic convergence theorem for the class of non-negative measurable functions. Its statement requires the following notation.

In analogy with the symbols $\nearrow$ and $\searrow$ used to describe increasing and decreasing sequences of sets, we shall write

$$f_n \nearrow f$$

whenever $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions that satisfies

$$f_n(x) \leq f_{n+1}(x) \text{ a.e } x, \text{ all } n \geq 1 \quad \text{and} \quad \lim_{n\to\infty} f_n(x) = f(x) \text{ a.e } x.$$

Similarly, we write $f_n \searrow f$ whenever

$$f_n(x) \geq f_{n+1}(x) \text{ a.e } x, \text{ all } n \geq 1 \quad \text{and} \quad \lim_{n\to\infty} f_n(x) = f(x) \text{ a.e } x.$$

Corollary 1.9 (Monotone convergence theorem) *Suppose $\{f_n\}$ is a sequence of non-negative measurable functions with $f_n \nearrow f$. Then*

$$\lim_{n\to\infty} \int f_n = \int f.$$

The monotone convergence theorem has the following useful consequence:

Corollary 1.10 *Consider a series $\sum_{k=1}^{\infty} a_k(x)$, where $a_k(x) \geq 0$ is measurable for every $k \geq 1$. Then*

$$\int \sum_{k=1}^{\infty} a_k(x)\, dx = \sum_{k=1}^{\infty} \int a_k(x)\, dx.$$

If $\sum_{k=1}^{\infty} \int a_k(x)\, dx$ is finite, then the series $\sum_{k=1}^{\infty} a_k(x)$ converges for a.e. x.

Proof. Let $f_n(x) = \sum_{k=1}^{n} a_k(x)$ and $f(x) = \sum_{k=1}^{\infty} a_k(x)$. The functions f_n are measurable, $f_n(x) \leq f_{n+1}(x)$, and $f_n(x) \to f(x)$ as n tends to infinity. Since

$$\int f_n = \sum_{k=1}^{n} \int a_k(x)\, dx,$$

the monotone convergence theorem implies

$$\sum_{k=1}^{\infty} \int a_k(x)\, dx = \int \sum_{k=1}^{\infty} a_k(x)\, dx.$$

If $\sum \int a_k < \infty$, then the above implies that $\sum_{k=1}^{\infty} a_k(x)$ is integrable, and by our earlier observation, we conclude that $\sum_{k=1}^{\infty} a_k(x)$ is finite almost everywhere.

We give two nice illustrations of this last corollary.

The first consists of another proof of the Borel-Cantelli lemma (see Exercise 16, Chapter 1), which says that if $E_1, E_2, \ldots$ is a collection of measurable subsets with $\sum m(E_k) < \infty$, then the set of points that belong to infinitely many sets E_k has measure zero. To prove this fact, we let

$$a_k(x) = \chi_{E_k}(x),$$

and note that a point x belongs to infinitely many sets E_k if and only if $\sum_{k=1}^{\infty} a_k(x) = \infty$. Our assumption on $\sum m(E_k)$ says precisely that $\sum_{k=1}^{\infty} \int a_k(x)\, dx < \infty$, and the corollary implies that $\sum_{k=1}^{\infty} a_k(x)$ is finite except possibly on a set of measure zero, and thus the Borel-Cantelli lemma is proved.

The second illustration will be useful in our discussion of approximations to the identity in Chapter 3. Consider the function

$$f(x) = \begin{cases} \frac{1}{|x|^{d+1}} & \text{if } x \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We prove that f is integrable outside any ball, $|x| \geq \epsilon$, and moreover

$$\int_{|x|\geq \epsilon} f(x)\, dx \leq \frac{C}{\epsilon}, \qquad \text{for some constant } C > 0.$$

Indeed, if we let $A_k = \{x \in \mathbb{R}^d : 2^k\epsilon < |x| \leq 2^{k+1}\epsilon\}$, and define

$$g(x) = \sum_{k=0}^{\infty} a_k(x) \quad \text{where} \quad a_k(x) = \frac{1}{(2^k\epsilon)^{d+1}} \chi_{A_k}(x),$$

then we must have $f(x) \leq g(x)$, and hence $\int f \leq \int g$. Since the set A_k is obtained from $\mathcal{A} = \{1 < |x| < 2\}$ by a dilation of factor $2^k\epsilon$, we have

by the relative dilation-invariance properties of the Lebesgue measure, that $m(A_k) = (2^k\epsilon)^d m(\mathcal{A})$. Also by Corollary 1.10, we see that

$$\int g = \sum_{k=0}^{\infty} \frac{m(A_k)}{(2^k\epsilon)^{d+1}} = m(\mathcal{A}) \sum_{k=0}^{\infty} \frac{(2^k\epsilon)^d}{(2^k\epsilon)^{d+1}} = \frac{C}{\epsilon},$$

where $C = 2m(\mathcal{A})$. Note that the same dilation-invariance property in fact shows that

$$\int_{|x|\geq\epsilon} \frac{dx}{|x|^{d+1}} = \frac{1}{\epsilon} \int_{|x|\geq 1} \frac{dx}{|x|^{d+1}}.$$

See also the identity (7) below.

Stage four: general case

If f is any real-valued measurable function on $\mathbb{R}^d$, we say that f is **Lebesgue integrable** (or just integrable) if the non-negative measurable function $|f|$ is integrable in the sense of the previous section.

If f is Lebesgue integrable, we give a meaning to its integral as follows. First, we may define

$$f^+(x) = \max(f(x), 0) \quad \text{and} \quad f^-(x) = \max(-f(x), 0),$$

so that both f^+ and f^- are non-negative and $f^+ - f^- = f$. Since $f^\pm \leq |f|$, both functions f^+ and f^- are integrable whenever f is, and we then define the **Lebesgue integral** of f by

$$\int f = \int f^+ - \int f^-.$$

In practice one encounters many decompositions $f = f_1 - f_2$, where f_1, f_2 are both non-negative integrable functions, and one would expect that regardless of the decomposition of f, we always have

$$\int f = \int f_1 - \int f_2.$$

In other words, the definition of the integral should be independent of the decomposition $f = f_1 - f_2$. To see why this is so, suppose $f = g_1 - g_2$ is another decomposition where both g_1 and g_2 are non-negative and integrable. Since $f_1 - f_2 = g_1 - g_2$ we have $f_1 + g_2 = g_1 + f_2$; but both

sides of this last identity consist of positive measurable functions, so the linearity of the integral in this case yields

$$\int f_1 + \int g_2 = \int g_1 + \int f_2.$$

Since all integrals involved are finite, we find the desired result

$$\int f_1 - \int f_2 = \int g_1 - \int g_2.$$

In considering the above definitions it is useful to keep in mind the following small observations. Both the integrability of f, and the value of its integral are unchanged if we modify f arbitrarily on a set of measure zero. It is therefore useful to adopt the convention that in the context of integration we allow our functions to be undefined on sets of measure zero. Moreover, if f is integrable, then by (v) of Proposition 1.6, it is finite-valued almost everywhere. Thus, availing ourselves of the above convention, we can always add two integrable functions f and g, since the ambiguity of $f + g$, due to the extended values of each, resides in a set of measure zero. Moreover, we note that when speaking of a function f, we are, in effect, also speaking about the collection of all functions that equal f almost everywhere.

Simple applications of the definition and the properties proved previously yield all the elementary properties of the integral:

Proposition 1.11 *The integral of Lebesgue integrable functions is linear, additive, monotonic, and satisfies the triangle inequality.*

We now gather two results which, although instructive in their own right, are also needed in the proof of the next theorem.

Proposition 1.12 *Suppose f is integrable on $\mathbb{R}^d$. Then for every $\epsilon > 0$:*

(i) *There exists a set of finite measure B (a ball, for example) such that*

$$\int_{B^c} |f| < \epsilon.$$

(ii) *There is a $\delta > 0$ such that*

$$\int_E |f| < \epsilon \qquad \textit{whenever } m(E) < \delta.$$

The last condition is known as absolute continuity.

Proof. By replacing f with $|f|$ we may assume without loss of generality that $f \geq 0$.

For the first part, let B_N denote the ball of radius N centered at the origin, and note that if $f_N(x) = f(x)\chi_{B_N}(x)$, then $f_N \geq 0$ is measurable, $f_N(x) \leq f_{N+1}(x)$, and $\lim_{N\to\infty} f_N(x) = f(x)$. By the monotone convergence theorem, we must have

$$\lim_{N\to\infty} \int f_N = \int f.$$

In particular, for some large N,

$$0 \leq \int f - \int f\chi_{B_N} < \epsilon,$$

and since $1 - \chi_{B_N} = \chi_{B_N^c}$, this implies $\int_{B_N^c} f < \epsilon$, as we set out to prove.

For the second part, assuming again that $f \geq 0$, we let $f_N(x) = f(x)\chi_{E_N}$ where

$$E_N = \{x : f(x) \leq N\}.$$

Once again, $f_N \geq 0$ is measurable, $f_N(x) \leq f_{N+1}(x)$, and given $\epsilon > 0$ there exists (by the monotone convergence theorem) an integer $N > 0$ such that

$$\int (f - f_N) < \frac{\epsilon}{2}.$$

We now pick $\delta > 0$ so that $N\delta < \epsilon/2$. If $m(E) < \delta$, then

$$\begin{aligned}
\int_E f &= \int_E (f - f_N) + \int_E f_N \\
&\leq \int (f - f_N) + \int_E f_N \\
&\leq \int (f - f_N) + Nm(E) \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

This concludes the proof of the proposition.

Intuitively, integrable functions should in some sense vanish at infinity since their integrals are finite, and the first part of the proposition attaches a precise meaning to this intuition. One should observe, however,

that integrability need not guarantee the more naive pointwise vanishing as $|x|$ becomes large. See Exercise 6.

We are now ready to prove a cornerstone of the theory of Lebesgue integration, the dominated convergence theorem. It can be viewed as a culmination of our efforts, and is a general statement about the interplay between limits and integrals.

Theorem 1.13 *Suppose $\{f_n\}$ is a sequence of measurable functions such that $f_n(x) \to f(x)$ a.e. x, as n tends to infinity. If $|f_n(x)| \leq g(x)$, where g is integrable, then*

$$\int |f_n - f| \to 0 \qquad \text{as } n \to \infty,$$

and consequently

$$\int f_n \to \int f \qquad \text{as } n \to \infty.$$

Proof. For each $N \geq 0$ let $E_N = \{x : |x| \leq N,\ g(x) \leq N\}$. Given $\epsilon > 0$, we may argue as in the first part of the previous lemma, to see that there exists N so that $\int_{E_N^c} g < \epsilon$. Then the functions $f_n \chi_{E_N}$ are bounded (by N) and supported on a set of finite measure, so that by the bounded convergence theorem, we have

$$\int_{E_N} |f_n - f| < \epsilon, \qquad \text{for all large } n.$$

Hence, we obtain the estimate

$$\begin{aligned}\int |f_n - f| &= \int_{E_N} |f_n - f| + \int_{E_N^c} |f_n - f| \\ &\leq \int_{E_N} |f_n - f| + 2\int_{E_N^c} g \\ &\leq \epsilon + 2\epsilon = 3\epsilon\end{aligned}$$

for all large n. This proves the theorem.

Complex-valued functions

If f is a complex-valued function on $\mathbb{R}^d$, we may write it as

$$f(x) = u(x) + iv(x),$$

where u and v are real-valued functions called the real and imaginary parts of f, respectively. The function f is measurable if and only if both u and v are measurable. We then say that f is **Lebesgue integrable** if the function $|f(x)| = (u(x)^2 + v(x)^2)^{1/2}$ (which is non-negative) is Lebesgue integrable in the sense defined previously.

It is clear that

$$|u(x)| \leq |f(x)| \quad \text{and} \quad |v(x)| \leq |f(x)|.$$

Also, if $a, b \geq 0$, one has $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$, so that

$$|f(x)| \leq |u(x)| + |v(x)|.$$

As a result of these simple inequalities, we deduce that a complex-valued function is integrable if and only if both its real and imaginary parts are integrable. Then, the Lebesgue integral of f is defined by

$$\int f(x)\,dx = \int u(x)\,dx + i \int v(x)\,dx.$$

Finally, if E is a measurable subset of $\mathbb{R}^d$, and f is a complex-valued measurable function on E, we say that f is Lebesgue integrable on E if $f\chi_E$ is integrable on $\mathbb{R}^d$, and we define $\int_E f = \int f\chi_E$.

The collection of all complex-valued integrable functions on a measurable subset $E \subset \mathbb{R}^d$ forms a vector space over $\mathbb{C}$. Indeed, if f and g are integrable, then so is $f+g$, since the triangle inequality gives $|(f+g)(x)| \leq |f(x)| + |g(x)|$, and monotonicity of the integral then yields

$$\int_E |f+g| \leq \int_E |f| + \int_E |g| < \infty.$$

Also, it is clear that if $a \in \mathbb{C}$ and if f is integrable, then so is af. Finally, the integral continues to be linear over $\mathbb{C}$.

2 The space L^1 of integrable functions

The fact that the integrable functions form a vector space is an important observation about the algebraic properties of such functions. A fundamental analytic fact is that this vector space is complete in the appropriate norm.

For any integrable function f on $\mathbb{R}^d$ we define the **norm**[2] of f,

$$\|f\| = \|f\|_{L^1} = \|f\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f(x)|\, dx.$$

The collection of all integrable functions with the above norm gives a (somewhat imprecise) definition of the space $L^1(\mathbb{R}^d)$. We also note that $\|f\| = 0$ if and only if $f = 0$ almost everywhere (see Proposition 1.6), and this simple property of the norm reflects the practice we have already adopted not to distinguish two functions that agree almost everywhere. With this in mind, we take the precise definition of $L^1(\mathbb{R}^d)$ to be the space of equivalence classes of integrable functions, where we define two functions to be **equivalent** if they agree almost everywhere. Often, however, it is convenient to retain the (imprecise) terminology that an element $f \in L^1(\mathbb{R}^d)$ is an integrable function, even though it is only an equivalence class of such functions. Note that by the above, the norm $\|f\|$ of an element $f \in L^1(\mathbb{R}^d)$ is well-defined by the choice of any integrable function in its equivalence class. Moreover, $L^1(\mathbb{R}^d)$ inherits the property that it is a vector space. This and other straightforward facts are summarized in the following proposition.

Proposition 2.1 *Suppose f and g are two functions in $L^1(\mathbb{R}^d)$.*

(i) $\|af\|_{L^1(\mathbb{R}^d)} = |a|\, \|f\|_{L^1(\mathbb{R}^d)}$ *for all* $a \in \mathbb{C}$.

(ii) $\|f + g\|_{L^1(\mathbb{R}^d)} \le \|f\|_{L^1(\mathbb{R}^d)} + \|g\|_{L^1(\mathbb{R}^d)}$.

(iii) $\|f\|_{L^1(\mathbb{R}^d)} = 0$ *if and only if* $f = 0$ *a.e.*

(iv) $d(f, g) = \|f - g\|_{L^1(\mathbb{R}^d)}$ *defines a metric on* $L^1(\mathbb{R}^d)$.

In (iv), we mean that d satisfies the following conditions. First, $d(f, g) \ge 0$ for all integrable functions f and g, and $d(f, g) = 0$ if and only if $f = g$ a.e. Also, $d(f, g) = d(g, f)$, and finally, d satisfies the triangle inequality

$$d(f, g) \le d(f, h) + d(h, g), \qquad \text{for all } f, g, h \in L^1(\mathbb{R}^d).$$

A space V with a metric d is said to be **complete** if for every Cauchy sequence $\{x_k\}$ in V (that is, $d(x_k, x_\ell) \to 0$ as $k, \ell \to \infty$) there exists $x \in V$ such that $\lim_{k\to\infty} x_k = x$ in the sense that

$$d(x_k, x) \to 0, \qquad \text{as } k \to \infty.$$

Our main goal of completing the space of Riemann integrable functions will be attained once we have established the next important theorem.

---

[2] In this chapter the only norm we consider is the L^1-norm, so we often write $\|f\|$ for $\|f\|_{L^1}$. Later, we shall have occasion to consider other norms, and then we shall modify our notation accordingly.

Theorem 2.2 (Riesz-Fischer) *The vector space L^1 is complete in its metric.*

Proof. Suppose $\{f_n\}$ is a Cauchy sequence in the norm, so that $\|f_n - f_m\| \to 0$ as $n, m \to \infty$. The plan of the proof is to extract a subsequence of $\{f_n\}$ that converges to f, both pointwise almost everywhere and in the norm.

Under ideal circumstances we would have that the sequence $\{f_n\}$ converges almost everywhere to a limit f, and we would then prove that the sequence converges to f also in the norm. Unfortunately, almost everywhere convergence does not hold for general Cauchy sequences (see Exercise 12). The main point, however, is that if the convergence in the norm is rapid enough, then almost everywhere convergence is a consequence, and this can be achieved by dealing with an appropriate subsequence of the original sequence.

Indeed, consider a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ of $\{f_n\}$ with the following property:

$$\|f_{n_{k+1}} - f_{n_k}\| \leq 2^{-k}, \qquad \text{for all } k \geq 1.$$

The existence of such a subsequence is guaranteed by the fact that $\|f_n - f_m\| \leq \epsilon$ whenever $n, m \geq N(\epsilon)$, so that it suffices to take $n_k = N(2^{-k})$.

We now consider the series whose convergence will be seen below,

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty}(f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty}|f_{n_{k+1}}(x) - f_{n_k}(x)|,$$

and note that

$$\int |f_{n_1}| + \sum_{k=1}^{\infty}\int |f_{n_{k+1}} - f_{n_k}| \leq \int |f_{n_1}| + \sum_{k=1}^{\infty} 2^{-k} < \infty.$$

So the monotone convergence theorem implies that g is integrable, and since $|f| \leq g$, hence so is f. In particular, the series defining f converges almost everywhere, and since the partial sums of this series are precisely the f_{n_k} (by construction of the telescopic series), we find that

$$f_{n_k}(x) \to f(x) \qquad \text{a.e. } x.$$

To prove that $f_{n_k} \to f$ in L^1 as well, we simply observe that $|f - f_{n_k}| \leq g$ for all k, and apply the dominated convergence theorem to get $\|f_{n_k} - f\|_{L^1} \to 0$ as k tends to infinity.

Finally, the last step of the proof consists in recalling that $\{f_n\}$ is Cauchy. Given ϵ, there exists N such that for all $n, m > N$ we have $\|f_n - f_m\| < \epsilon/2$. If n_k is chosen so that $n_k > N$, and $\|f_{n_k} - f\| < \epsilon/2$, then the triangle inequality implies

$$\|f_n - f\| \leq \|f_n - f_{n_k}\| + \|f_{n_k} - f\| < \epsilon$$

whenever $n > N$. Thus $\{f_n\}$ has the limit f in L^1, and the proof of the theorem is complete.

Since every sequence that converges in the norm is a Cauchy sequence in that norm, the argument in the proof of the theorem yields the following.

Corollary 2.3 *If $\{f_n\}_{n=1}^{\infty}$ converges to f in L^1, then there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ such that*

$$f_{n_k}(x) \to f(x) \qquad a.e.\ x.$$

We say that a family $\mathcal{G}$ of integrable functions is **dense** in L^1 if for any $f \in L^1$ and $\epsilon > 0$, there exists $g \in \mathcal{G}$ so that $\|f - g\|_{L^1} < \epsilon$. Fortunately we are familiar with many families that are dense in L^1, and we describe some in the theorem that follows. These are useful when one is faced with the problem of proving some fact or identity involving integrable functions. In this situation a general principle applies: the result is often easier to prove for a more restrictive class of functions (like the ones in the theorem below), and then a density (or limiting) argument yields the result in general.

Theorem 2.4 *The following families of functions are dense in $L^1(\mathbb{R}^d)$:*

(i) *The simple functions.*

(ii) *The step functions.*

(iii) *The continuous functions of compact support.*

Proof. Let f be an integrable function on $\mathbb{R}^d$. First, we may assume that f is real-valued, because we may approximate its real and imaginary parts independently. If this is the case, we may then write $f = f^+ - f^-$, where $f^+, f^- \geq 0$, and it now suffices to prove the theorem when $f \geq 0$.

For (i), Theorem 4.1 in Chapter 1 guarantees the existence of a sequence $\{\varphi_k\}$ of non-negative simple functions that increase to f pointwise. By the dominated convergence theorem (or even simply the monotone convergence theorem) we then have

$$\|f - \varphi_k\|_{L^1} \to 0 \quad \text{as } k \to \infty.$$

Thus there are simple functions that are arbitrarily close to f in the L^1 norm.

For (ii), we first note that by (i) it suffices to approximate simple functions by step functions. Then, we recall that a simple function is a finite linear combination of characteristic functions of sets of finite measure, so it suffices to show that if E is such a set, then there is a step function ψ so that $\|\chi_E - \psi\|_{L^1}$ is small. However, we now recall that this argument was already carried out in the proof of Theorem 4.3, Chapter 1. Indeed, there it is shown that there is an almost disjoint family of rectangles $\{R_j\}$ with $m(E \triangle \bigcup_{j=1}^M R_j) \leq 2\epsilon$. Thus χ_E and $\psi = \sum_j \chi_{R_j}$ differ at most on a set of measure 2ϵ, and as a result we find that $\|\chi_E - \psi\|_{L^1} < 2\epsilon$.

By (ii), it suffices to establish (iii) when f is the characteristic function of a rectangle. In the one-dimensional case, where f is the characteristic function of an interval $[a, b]$, we may choose a continuous piecewise linear function g defined by

$$g(x) = \begin{cases} 1 & \text{if } a \leq x \leq b, \\ 0 & \text{if } x \leq a - \epsilon \text{ or } x \geq b + \epsilon, \end{cases}$$

and with g linear on the intervals $[a - \epsilon, a]$ and $[b, b + \epsilon]$. Then $\|f - g\|_{L^1} < 2\epsilon$. In d dimensions, it suffices to note that the characteristic function of a rectangle is the product of characteristic functions of intervals. Then, the desired continuous function of compact support is simply the product of functions like g defined above.

The results above for $L^1(\mathbb{R}^d)$ lead immediately to an extension in which $\mathbb{R}^d$ can be replaced by any fixed subset E of positive measure. In fact if E is such a subset, we can define $L^1(E)$ and carry out the arguments that are analogous to $L^1(\mathbb{R}^d)$. Better yet, we can proceed by extending any function f on E by setting $\tilde{f} = f$ on E and $\tilde{f} = 0$ on E^c, and defining $\|f\|_{L^1(E)} = \|\tilde{f}\|_{L^1(\mathbb{R}^d)}$. The analogues of Proposition 2.1 and Theorem 2.2 then hold for the space $L^1(E)$.

Invariance Properties

If f is a function defined on $\mathbb{R}^d$, the **translation** of f by a vector $h \in \mathbb{R}^d$ is the function f_h, defined by $f_h(x) = f(x-h)$. Here we want to examine some basic aspects of translations of integrable functions.

First, there is the translation-invariance of the integral. One way to state this is as follows: if f is an integrable function, then so is f_h and

$$\int_{\mathbb{R}^d} f(x-h)\,dx = \int_{\mathbb{R}^d} f(x)\,dx. \tag{4}$$

We check this assertion first when $f = \chi_E$, the characteristic function of a measurable set E. Then obviously $f_h = \chi_{E_h}$, where $E_h = \{x+h : x \in E\}$, and thus the assertion follows because $m(E_h) = m(E)$ (see Section 3 in Chapter 1). As a result of linearity, the identity (4) holds for all simple functions. Now if f is non-negative and $\{\varphi_n\}$ is a sequence of simple functions that increase pointwise a.e to f (such a sequence exists by Theorem 4.1 in the previous chapter), then $\{(\varphi_n)_h\}$ is a sequence of simple functions that increase to f_h pointwise a.e, and the monotone convergence theorem implies (4) in this special case. Thus, if f is complex-valued and integrable we see that $\int_{\mathbb{R}^d} |f(x-h)|\,dx = \int_{\mathbb{R}^d} |f(x)|\,dx$, which shows that $f_h \in L^1(\mathbb{R}^d)$ and also $\|f_h\| = \|f\|$. From the definitions, we then conclude that (4) holds whenever $f \in L^1$.

Incidentally, using the relative invariance of Lebesgue measure under dilations and reflections (Section 3, Chapter 1) one can prove in the same way that if $f(x)$ is integrable, so is $f(\delta x)$, $\delta > 0$, and $f(-x)$, and

$$\delta^d \int_{\mathbb{R}^d} f(\delta x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx, \quad \text{while} \quad \int_{\mathbb{R}^d} f(-x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx. \tag{5}$$

We digress to record for later use two useful consequences of the above invariance properties:

(i) Suppose that f and g are a pair of measurable functions on $\mathbb{R}^d$ so that for some fixed $x \in \mathbb{R}^d$ the function $y \mapsto f(x-y)g(y)$ is integrable. As a consequence, the function $y \mapsto f(y)g(x-y)$ is then also integrable and we have

$$\int_{\mathbb{R}^d} f(x-y)g(y)\,dy = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy. \tag{6}$$

This follows from (4) and (5) on making the change of variables which replaces y by $x-y$, and noting that this change is a combination of a translation and a reflection.

The integral on the left-hand side is denoted by $(f * g)(x)$ and is defined as the **convolution** of f and g. Thus (6) asserts the commutativity of the convolution product.

(ii) Using (5) one has that for all $\epsilon > 0$

$$\int_{|x|\geq \epsilon} \frac{dx}{|x|^a} = \epsilon^{-a+d} \int_{|x|\geq 1} \frac{dx}{|x|^a} \quad \text{whenever } a > d, \tag{7}$$

and

$$\int_{|x|\leq \epsilon} \frac{dx}{|x|^a} = \epsilon^{-a+d} \int_{|x|\leq 1} \frac{dx}{|x|^a} \quad \text{whenever } a < d. \tag{8}$$

It can also be seen that the integrals $\int_{|x|\geq 1} \frac{dx}{|x|^a}$ and $\int_{|x|\leq 1} \frac{dx}{|x|^a}$ (respectively, when $a > d$ and $a < d$) are finite by the argument that appears after Corollary 1.10.

Translations and continuity

We shall next examine how continuity properties of f are related to the way the translations f_h vary with h. Note that for any given $x \in \mathbb{R}^d$, the statement that $f_h(x) \to f(x)$ as $h \to 0$ is the same as the continuity of f at the point x.

However, a general f which is integrable may be discontinuous at every x, even when corrected on a set of measure zero; see Exercise 15. Nevertheless, there is an overall continuity that an arbitrary $f \in L^1(\mathbb{R}^d)$ enjoys, one that holds in the norm.

Proposition 2.5 *Suppose $f \in L^1(\mathbb{R}^d)$. Then*

$$\|f_h - f\|_{L^1} \to 0 \qquad \text{as } h \to 0.$$

The proof is a simple consequence of the approximation of integrable functions by continuous functions of compact support as given in Theorem 2.4. In fact for any $\epsilon > 0$, we can find such a function g so that $\|f - g\| < \epsilon$. Now

$$f_h - f = (g_h - g) + (f_h - g_h) - (f - g).$$

However, $\|f_h - g_h\| = \|f - g\| < \epsilon$, while since g is continuous and has compact support we have that clearly

$$\|g_h - g\| = \int_{\mathbb{R}^d} |g(x - h) - g(x)|\, dx \to 0 \qquad \text{as } h \to 0.$$

So if $|h| < \delta$, where δ is sufficiently small, then $\|g_h - g\| < \epsilon$, and as a result $\|f_h - f\| < 3\epsilon$, whenever $|h| < \delta$.

3 Fubini's theorem

In elementary calculus integrals of continuous functions of several variables are often calculated by iterating one-dimensional integrals. We shall now examine this important analytic device from the general point of view of Lebesgue integration in $\mathbb{R}^d$, and we shall see that a number of interesting issues arise.

In general, we may write $\mathbb{R}^d$ as a product

$$\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \qquad \text{where } d = d_1 + d_2, \text{ and } d_1, d_2 \geq 1.$$

A point in $\mathbb{R}^d$ then takes the form (x, y), where $x \in \mathbb{R}^{d_1}$ and $y \in \mathbb{R}^{d_2}$. With such a decomposition of $\mathbb{R}^d$ in mind, the general notion of a slice, formed by fixing one variable, becomes natural. If f is a function in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, the **slice** of f corresponding to $y \in \mathbb{R}^{d_2}$ is the function f^y of the $x \in \mathbb{R}^{d_1}$ variable, given by

$$f^y(x) = f(x, y).$$

Similarly, the slice of f for a fixed $x \in \mathbb{R}^{d_1}$ is $f_x(y) = f(x, y)$.

In the case of a set $E \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ we define its **slices** by

$$E^y = \{x \in \mathbb{R}^{d_1} : \ (x, y) \in E\} \quad \text{and} \quad E_x = \{y \in \mathbb{R}^{d_2} : \ (x, y) \in E\}.$$

See Figure 1 for an illustration.

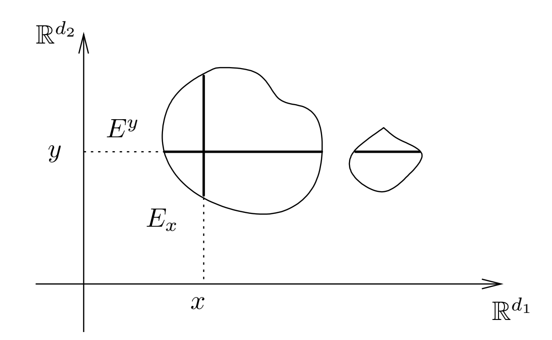


Figure 1. Slices E^y and E_x (for fixed x and y) of a set E

3.1 Statement and proof of the theorem

That the theorem that follows is not entirely straightforward is clear from the first difficulty that arises in its formulation, involving the measurability of the functions and sets in question. In fact, even with the

assumption that f is measurable on $\mathbb{R}^d$, it is not necessarily true that the slice f^y is measurable on $\mathbb{R}^{d_1}$ for each y; nor does the corresponding assertion necessarily hold for a measurable set: the slice E^y may not be measurable for each y. An easy example arises in $\mathbb{R}^2$ by placing a one-dimensional non-measurable set on the x-axis; the set E in $\mathbb{R}^2$ has measure zero, but E^y is not measurable for $y = 0$. What saves us is that, nevertheless, measurability holds for almost all slices.

The main theorem is as follows. We recall that by definition all integrable functions are measurable.

Theorem 3.1 *Suppose $f(x, y)$ is integrable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for almost every $y \in \mathbb{R}^{d_2}$:*

(i) *The slice f^y is integrable on $\mathbb{R}^{d_1}$.*

(ii) *The function defined by $\int_{\mathbb{R}^{d_1}} f^y(x)\, dx$ is integrable on $\mathbb{R}^{d_2}$.*

Moreover:

(iii) $$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x, y)\, dx \right) dy = \int_{\mathbb{R}^d} f.$$

Clearly, the theorem is symmetric in x and y so that we also may conclude that the slice f_x is integrable on $\mathbb{R}^{d_2}$ for a.e. x. Moreover, $\int_{\mathbb{R}^{d_2}} f_x(y)\, dy$ is integrable, and

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x, y)\, dy \right) dx = \int_{\mathbb{R}^d} f.$$

In particular, Fubini's theorem states that the integral of f on $\mathbb{R}^d$ can be computed by iterating lower-dimensional integrals, and that the iterations can be taken in any order

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x, y)\, dx \right) dy = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x, y)\, dy \right) dx = \int_{\mathbb{R}^d} f.$$

We first note that we may assume that f is real-valued, since the theorem then applies to the real and imaginary parts of a complex-valued function. The proof of Fubini's theorem which we give next consists of a sequence of six steps. We begin by letting $\mathcal{F}$ denote the set of integrable functions on $\mathbb{R}^d$ which satisfy all three conclusions in the theorem, and set out to prove that $L^1(\mathbb{R}^d) \subset \mathcal{F}$.

We proceed by first showing that $\mathcal{F}$ is closed under operations such as linear combinations (Step 1) and limits (Step 2). Then we begin to

construct families of functions in $\mathcal{F}$. Since any integrable function is the "limit" of simple functions, and simple functions are themselves linear combinations of sets of finite measure, the goal quickly becomes to prove that χ_E belongs to $\mathcal{F}$ whenever E is a measurable subset of $\mathbb{R}^d$ with finite measure. To achieve this goal, we begin with rectangles and work our way up to sets of type G_δ (Step 3), and sets of measure zero (Step 4). Finally, a limiting argument shows that all integrable functions are in $\mathcal{F}$. This will complete the proof of Fubini's theorem.

Step 1. Any finite linear combination of functions in $\mathcal{F}$ also belongs to $\mathcal{F}$.

Indeed, let $\{f_k\}_{k=1}^N \subset \mathcal{F}$. For each k there exists a set $A_k \subset \mathbb{R}^{d_2}$ of measure 0 so that f_k^y is integrable on $\mathbb{R}^{d_1}$ whenever $y \notin A_k$. Then, if $A = \bigcup_{k=1}^N A_k$, the set A has measure 0, and in the complement of A, the y-slice corresponding to any finite linear combination of the f_k is measurable, and also integrable. By linearity of the integral, we then conclude that any linear combination of the f_k's belongs to $\mathcal{F}$.

Step 2. Suppose $\{f_k\}$ is a sequence of measurable functions in $\mathcal{F}$ so that $f_k \nearrow f$ or $f_k \searrow f$, where f is integrable (on $\mathbb{R}^d$). Then $f \in \mathcal{F}$.

By taking $-f_k$ instead of f_k if necessary, we note that it suffices to consider the case of an increasing sequence. Also, we may replace f_k by $f_k - f_1$ and assume that the f_k's are non-negative. Now, we observe that an application of the monotone convergence theorem (Corollary 1.9) yields

$$\lim_{k\to\infty} \int_{\mathbb{R}^d} f_k(x,y)\,dx\,dy = \int_{\mathbb{R}^d} f(x,y)\,dx\,dy. \tag{9}$$

By assumption, for each k there exists a set $A_k \subset \mathbb{R}^{d_2}$, so that f_k^y is integrable on $\mathbb{R}^{d_1}$ whenever $y \notin A_k$. If $A = \bigcup_{k=1}^\infty A_k$, then $m(A) = 0$ in $\mathbb{R}^{d_2}$, and if $y \notin A$, then f_k^y is integrable on $\mathbb{R}^{d_1}$ for all k, and, by the monotone convergence theorem, we find that

$$g_k(y) = \int_{\mathbb{R}^{d_1}} f_k^y(x)\,dx \quad \text{increases to a limit} \quad g(y) = \int_{\mathbb{R}^{d_1}} f^y(x)\,dx$$

as k tends to infinity. By assumption, each $g_k(y)$ is integrable, so that another application of the monotone convergence theorem yields

$$\int_{\mathbb{R}^{d_2}} g_k(y)\,dy \to \int_{\mathbb{R}^{d_2}} g(y)\,dy \quad \text{as } k \to \infty. \tag{10}$$

By the assumption that $f_k \in \mathcal{F}$ we have

$$\int_{\mathbb{R}^{d_2}} g_k(y)\, dy = \int_{\mathbb{R}^d} f_k(x,y)\, dx\, dy,$$

and combining this fact with (9) and (10), we conclude that

$$\int_{\mathbb{R}^{d_2}} g(y)\, dy = \int_{\mathbb{R}^d} f(x,y)\, dx\, dy.$$

Since f is integrable, the right-hand integral is finite, and this proves that g is integrable. Consequently $g(y) < \infty$ a.e. y, hence f^y is integrable for a.e. y, and

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)\, dx \right) dy = \int_{\mathbb{R}^d} f(x,y)\, dx\, dy.$$

This proves that $f \in \mathcal{F}$ as desired.

Step 3. Any characteristic function of a set E that is a G_δ and of finite measure belongs to $\mathcal{F}$.

We proceed in stages of increasing order of generality.

(a) First suppose E is a bounded open cube in $\mathbb{R}^d$, such that $E = Q_1 \times Q_2$, where Q_1 and Q_2 are open cubes in $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, respectively. Then, for each y the function $\chi_E(x,y)$ is measurable in x, and integrable with

$$g(y) = \int_{\mathbb{R}^{d_1}} \chi_E(x,y)\, dx \begin{cases} |Q_1| & \text{if } y \in Q_2, \\ 0 & \text{otherwise.} \end{cases}$$

Consequently, $g = |Q_1|\chi_{Q_2}$ is also measurable and integrable, with

$$\int_{\mathbb{R}^{d_2}} g(y)\, dy = |Q_1|\,|Q_2|.$$

Since we initially have $\int_{\mathbb{R}^d} \chi_E(x,y)\, dx\, dy = |E| = |Q_1|\,|Q_2|$, we deduce that $\chi_E \in \mathcal{F}$.

(b) Now suppose E is a subset of the boundary of some closed cube. Then, since the boundary of a cube has measure 0 in $\mathbb{R}^d$, we have $\int_{\mathbb{R}^d} \chi_E(x,y)\, dx\, dy = 0$.

Next, we note, after an investigation of the various possibilities, that for almost every y, the slice E^y has measure 0 in $\mathbb{R}^{d_1}$, and therefore if $g(y) = \int_{\mathbb{R}^{d_1}} \chi_E(x,y)\, dx$ we have $g(y) = 0$ for a.e. y. As a consequence, $\int_{\mathbb{R}^{d_2}} g(y)\, dy = 0$, and therefore $\chi_E \in \mathcal{F}$.

(c) Suppose now E is a finite union of closed cubes whose interiors are disjoint, $E = \bigcup_{k=1}^{K} Q_k$. Then, if $\tilde{Q}_k$ denotes the interior of Q_k, we may write χ_E as a linear combination of the $\chi_{\tilde{Q}_k}$ and χ_{A_k} where A_k is a subset of the boundary of Q_k for $k = 1, \ldots, K$. By our previous analysis, we know that χ_{Q_k} and χ_{A_k} belong to $\mathcal{F}$ for all k, and since Step 1 guarantees that $\mathcal{F}$ is closed under finite linear combinations, we conclude that $\chi_E \in \mathcal{F}$, as desired.

(d) Next, we prove that if E is open and of finite measure, then $\chi_E \in \mathcal{F}$. This follows from taking a limit in the previous case. Indeed, by Theorem 1.4 in Chapter 1, we may write E as a countable union of almost disjoint closed cubes

$$E = \bigcup_{j=1}^{\infty} Q_j.$$

Consequently, if we let $f_k = \sum_{j=1}^{k} \chi_{Q_j}$, then we note that the functions f_k increase to $f = \chi_E$, which is integrable since $m(E)$ is finite. Therefore, we may conclude by Step 2 that $f \in \mathcal{F}$.

(e) Finally, if E is a G_δ of finite measure, then $\chi_E \in \mathcal{F}$. Indeed, by definition, there exist open sets $\tilde{\mathcal{O}}_1, \tilde{\mathcal{O}}_2, \ldots$, such that

$$E = \bigcap_{k=1}^{\infty} \tilde{\mathcal{O}}_k.$$

Since E has finite measure, there exists an open set $\tilde{\mathcal{O}}_0$ of finite measure with $E \subset \tilde{\mathcal{O}}_0$. If we let

$$\mathcal{O}_k = \mathcal{O}_0 \cap \bigcap_{j=1}^{k} \tilde{\mathcal{O}}_j,$$

then we note that we have a decreasing sequence of open sets of finite measure $\mathcal{O}_1 \supset \mathcal{O}_2 \supset \cdots$ with

$$E = \bigcap_{k=1}^{\infty} \mathcal{O}_k.$$

Therefore, the sequence of functions $f_k = \chi_{\mathcal{O}_k}$ decreases to $f = \chi_E$, and since $\chi_{\mathcal{O}_k} \in \mathcal{F}$ for all k by (d) above, we conclude by Step 2 that χ_E belongs to $\mathcal{F}$.

Step 4. If E has measure 0, then χ_E belongs to $\mathcal{F}$.

Indeed, since E is measurable, we may choose a set G of type G_δ with $E \subset G$ and $m(G) = 0$ (Corollary 3.5, Chapter 1). Since $\chi_G \in \mathcal{F}$ (by the previous step) we find that

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} \chi_G(x, y)\, dx \right) dy = \int_{\mathbb{R}^d} \chi_G = 0.$$

Therefore

$$\int_{\mathbb{R}^{d_1}} \chi_G(x, y)\, dx = 0 \qquad \text{for a.e. } y.$$

Consequently, the slice G^y has measure 0 for a.e. y. The simple observation that $E^y \subset G^y$ then shows that E^y has measure 0 for a.e. y, and $\int_{\mathbb{R}^{d_1}} \chi_E(x, y)\, dx = 0$ for a.e. y. Therefore,

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} \chi_E(x, y)\, dx \right) dy = 0 = \int_{\mathbb{R}^d} \chi_E,$$

and thus $\chi_E \in \mathcal{F}$, as was to be shown.

Step 5. If E is any measurable subset of $\mathbb{R}^d$ with finite measure, then χ_E belongs to $\mathcal{F}$.

To prove this, recall first that there exists a set of finite measure G of type G_δ, with $E \subset G$ and $m(G - E) = 0$. Since

$$\chi_E = \chi_G - \chi_{G-E},$$

and $\mathcal{F}$ is closed under linear combinations, we find that $\chi_E \in \mathcal{F}$, as desired.

Step 6. This is the final step, which consists of proving that if f is integrable, then $f \in \mathcal{F}$.

We note first that f has the decomposition $f = f^+ - f^-$, where both f^+ and f^- are non-negative and integrable, so by Step 1 we may assume that f is itself non-negative. By Theorem 4.1 in the previous chapter, there exists a sequence $\{\varphi_k\}$ of simple functions that increase to f. Since each φ_k is a finite linear combination of characteristic functions of sets with finite measure, we have $\varphi_k \in \mathcal{F}$ by Steps 5 and 1, hence $f \in \mathcal{F}$ by Step 2.

3.2 Applications of Fubini's theorem

Theorem 3.2 *Suppose $f(x, y)$ is a non-negative measurable function on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for almost every $y \in \mathbb{R}^{d_2}$:*

(i) *The slice* f^y *is measurable on* $\mathbb{R}^{d_1}$.

(ii) *The function defined by* $\int_{\mathbb{R}^{d_1}} f^y(x)\,dx$ *is measurable on* $\mathbb{R}^{d_2}$.

Moreover:

(iii) $\displaystyle\int_{\mathbb{R}^{d_2}}\left(\int_{\mathbb{R}^{d_1}} f(x,y)\,dx\right)dy = \int_{\mathbb{R}^d} f(x,y)\,dx\,dy$ *in the extended sense.*

In practice, this theorem is often used in conjunction with Fubini's theorem.[3] Indeed, suppose we are given a measurable function f on $\mathbb{R}^d$ and asked to compute $\int_{\mathbb{R}^d} f$. To justify the use of iterated integration, we first apply the present theorem to $|f|$. Using it, we may freely compute (or estimate) the iterated integrals of the non-negative function $|f|$. If these are finite, Theorem 3.2 guarantees that f is integrable, that is, $\int |f| < \infty$. Then the hypothesis in Fubini's theorem is verified, and we may use that theorem in the calculation of the integral of f.

Proof of Theorem 3.2. Consider the truncations

$$f_k(x,y) = \begin{cases} f(x,y) & \text{if } |(x,y)| < k \text{ and } f(x,y) < k, \\ 0 & \text{otherwise.} \end{cases}$$

Each f_k is integrable, and by part (i) in Fubini's theorem there exists a set $E_k \subset \mathbb{R}^{d_2}$ of measure 0 such that the slice $f_k^y(x)$ is measurable for all $y \in E_k^c$. Then, if we set $E = \bigcup_k E_k$, we find that $f^y(x)$ is measurable for all $y \in E^c$ and all k. Moreover, $m(E) = 0$. Since $f_k^y \nearrow f^y$, the monotone convergence theorem implies that if $y \notin E$, then

$$\int_{\mathbb{R}^{d_1}} f_k(x,y)\,dx \nearrow \int_{\mathbb{R}^{d_1}} f(x,y)\,dx \qquad \text{as } k \to \infty.$$

Again by Fubini's theorem, $\int_{\mathbb{R}^{d_1}} f_k(x,y)\,dx$ is measurable for all $y \in E^c$, hence so is $\int_{\mathbb{R}^{d_1}} f(x,y)\,dx$. Another application of the monotone convergence theorem then gives

$$\int_{\mathbb{R}^{d_2}}\left(\int_{\mathbb{R}^{d_1}} f_k(x,y)\,dx\right)dy \to \int_{\mathbb{R}^{d_2}}\left(\int_{\mathbb{R}^{d_1}} f(x,y)\,dx\right)dy. \tag{11}$$

By part (iii) in Fubini's theorem we know that

$$\int_{\mathbb{R}^{d_2}}\left(\int_{\mathbb{R}^{d_1}} f_k(x,y)\,dx\right)dy = \int_{\mathbb{R}^d} f_k. \tag{12}$$

[3]Theorem 3.2 was formulated by Tonelli. We will, however, use the short-hand of referring to it, as well as Theorem 3.1 and Corollary 3.3, as Fubini's theorem.

A final application of the monotone convergence theorem directly to f_k also gives

$$\int_{\mathbb{R}^d} f_k \to \int_{\mathbb{R}^d} f. \tag{13}$$

Combining (11), (12), and (13) completes the proof of Theorem 3.2.

Corollary 3.3 *If E is a measurable set in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, then for almost every $y \in \mathbb{R}^{d_2}$ the slice*

$$E^y = \{x \in \mathbb{R}^{d_1} : \ (x,y) \in E\}$$

is a measurable subset of $\mathbb{R}^{d_1}$. Moreover $m(E^y)$ is a measurable function of y and

$$m(E) = \int_{\mathbb{R}^{d_2}} m(E^y)\, dy.$$

This is an immediate consequence of the first part of Theorem 3.2 applied to the function χ_E. Clearly a symmetric result holds for the x-slices in $\mathbb{R}^{d_2}$.

We have thus established the basic fact that if E is measurable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, then for almost every $y \in \mathbb{R}^{d_2}$ the slice E^y is measurable in $\mathbb{R}^{d_1}$ (and also the symmetric statement with the roles of x and y interchanged). One might be tempted to think that the converse assertion holds. To see that this is not the case, note that if we let $\mathcal{N}$ denote a non-measurable subset of $\mathbb{R}$, and then define

$$E = [0,1] \times \mathcal{N} \subset \mathbb{R} \times \mathbb{R},$$

we see that

$$E^y = \begin{cases} [0,1] & \text{if } y \in \mathcal{N}, \\ \emptyset & \text{if } y \notin \mathcal{N}. \end{cases}$$

Thus E^y is measurable for every y. However, if E were measurable, then the corollary would imply that $E_x = \{y \in \mathbb{R} : \ (x,y) \in E\}$ is measurable for almost every $x \in \mathbb{R}$, which is not true since E_x is equal to $\mathcal{N}$ for all $x \in [0,1]$.

A more striking example is that of a set E in the unit square $[0,1] \times [0,1]$ that is not measurable, and yet the slices E^y and E_x are measurable with $m(E^y) = 0$ and $m(E_x) = 1$ for each $x, y \in [0,1]$. The construction of E is based on the existence of a highly paradoxical ordering $\prec$ of

the reals, with the property that $\{x : x \prec y\}$ is a countable set for each $y \in \mathbb{R}$. (The construction of this ordering is discussed in Problem 5.) Given this ordering we let

$$E = \{(x, y) \in [0, 1] \times [0, 1], \text{ with } x \prec y\}.$$

Note that for each $y \in [0, 1]$, $E^y = \{x : x \prec y\}$; thus E^y is countable and $m(E^y) = 0$. Similarly $m(E_x) = 1$, because E_x is the complement of a denumerable set in $[0, 1]$. If E were measurable, it would contradict the formula in Corollary 3.3.

In relating a set E to its slices E_x and E^y, matters are straightforward for the basic sets which arise when we consider $\mathbb{R}^d$ as the product $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. These are the **product sets** $E = E_1 \times E_2$, where $E_j \subset \mathbb{R}^{d_j}$.

Proposition 3.4 *If $E = E_1 \times E_2$ is a measurable subset of $\mathbb{R}^d$, and $m_*(E_2) > 0$, then E_1 is measurable.*

Proof. By Corollary 3.3, we know that for a.e. $y \in \mathbb{R}^{d_2}$, the slice function

$$(\chi_{E_1 \times E_2})^y(x) = \chi_{E_1}(x)\chi_{E_2}(y)$$

is measurable as a function of x. In fact, we claim that there is some $y \in E_2$ such that the above slice function is measurable in x; for such a y we would have $\chi_{E_1 \times E_2}(x, y) = \chi_{E_1}(x)$, and this would imply that E_1 is measurable.

To prove the existence of such a y, we use the assumption that $m_*(E_2) > 0$. Indeed, let F denote the set of $y \in \mathbb{R}^{d_2}$ such that the slice E^y is measurable. Then $m(F^c) = 0$ (by the previous corollary). However, $E_2 \cap F$ is not empty because $m_*(E_2 \cap F) > 0$. To see this, note that $E_2 = (E_2 \cap F) \bigcup (E_2 \cap F^c)$, hence

$$0 < m_*(E_2) \leq m_*(E_2 \cap F) + m_*(E_2 \cap F^c) = m_*(E_2 \cap F),$$

because $E_2 \cap F^c$ is a subset of a set of measure zero.

To deal with a converse of the above result, we need the following lemma.

Lemma 3.5 *If $E_1 \subset \mathbb{R}^{d_1}$ and $E_2 \subset \mathbb{R}^{d_2}$, then*

$$m_*(E_1 \times E_2) \leq m_*(E_1)\, m_*(E_2),$$

with the understanding that if one of the sets E_j has exterior measure zero, then $m_*(E_1 \times E_2) = 0$.

Proof. Let $\epsilon > 0$. By definition, we can find cubes $\{Q_k\}_{k=1}^{\infty}$ in $\mathbb{R}^{d_1}$ and $\{Q'_\ell\}_{\ell=1}^{\infty}$ in $\mathbb{R}^{d_2}$ such that

$$E_1 \subset \bigcup_{k=1}^{\infty} Q_k, \quad \text{and} \quad E_2 \subset \bigcup_{\ell=1}^{\infty} Q'_\ell$$

and

$$\sum_{k=1}^{\infty} |Q_k| \leq m_*(E_1) + \epsilon \quad \text{and} \quad \sum_{\ell=1}^{\infty} |Q'_\ell| \leq m_*(E_2) + \epsilon.$$

Since $E_1 \times E_2 \subset \bigcup_{k,\ell=1}^{\infty} Q_k \times Q'_\ell$, the sub-additivity of the exterior measure yields

$$\begin{aligned} m_*(E_1 \times E_2) &\leq \sum_{k,\ell=1}^{\infty} |Q_k \times Q'_\ell| \\ &= \left(\sum_{k=1}^{\infty} |Q_k|\right)\left(\sum_{\ell=1}^{\infty} |Q'_\ell|\right) \\ &\leq (m_*(E_1) + \epsilon)(m_*(E_2) + \epsilon). \end{aligned}$$

If neither E_1 nor E_2 has exterior measure 0, then from the above we find

$$m_*(E_1 \times E_2) \leq m_*(E_1)\, m_*(E_2) + O(\epsilon),$$

and since ϵ is arbitrary, we must have $m_*(E_1 \times E_2) \leq m_*(E_1)\, m_*(E_2)$.

If for instance $m_*(E_1) = 0$, consider for each positive integer j the set $E_2^j = E_2 \cap \{y \in \mathbb{R}^{d_2} : |y| \leq j\}$. Then, by the above argument, we find that $m_*(E_1 \times E_2^j) = 0$. Since $(E_1 \times E_2^j) \nearrow (E_1 \times E_2)$ as $j \to \infty$, we conclude that $m_*(E_1 \times E_2) = 0$.

Proposition 3.6 *Suppose E_1 and E_2 are measurable subsets of $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, respectively. Then $E = E_1 \times E_2$ is a measurable subset of $\mathbb{R}^d$. Moreover,*

$$m(E) = m(E_1)\, m(E_2),$$

with the understanding that if one of the sets E_j has measure zero, then $m(E) = 0$.

Proof. It suffices to prove that E is measurable, because then the assertion about $m(E)$ follows from Corollary 3.3. Since each set E_j is

measurable, there exist sets $G_j \subset \mathbb{R}^{d_j}$ of type G_δ, with $G_j \supset E_j$ and $m_*(G_j - E_j) = 0$ for each $j = 1, 2$. (See Corollary 3.5 in Chapter 1.) Clearly, $G = G_1 \times G_2$ is measurable in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and

$$(G_1 \times G_2) - (E_1 \times E_2) \subset ((G_1 - E_1) \times G_2) \cup (G_1 \times (G_2 - E_2)).$$

By the lemma we conclude that $m_*(G - E) = 0$, hence E is measurable.

As a consequence of this proposition we have the following.

Corollary 3.7 *Suppose f is a measurable function on $\mathbb{R}^{d_1}$. Then the function $\tilde{f}$ defined by $\tilde{f}(x, y) = f(x)$ is measurable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$.*

Proof. To see this, we may assume that f is real-valued, and recall first that if $a \in \mathbb{R}$ and $E_1 = \{x \in \mathbb{R}^{d_1} : f(x) < a\}$, then E_1 is measurable by definition. Since

$$\{(x, y) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} : \ \tilde{f}(x, y) < a\} = E_1 \times \mathbb{R}^{d_2},$$

the previous proposition shows that $\{\tilde{f}(x, y) < a\}$ is measurable for each $a \in \mathbb{R}$. Thus $\tilde{f}(x, y)$ is a measurable function on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, as desired.

Finally, we return to an interpretation of the integral that arose first in the calculus. We have in mind the notion that $\int f$ describes the "area" under the graph of f. Here we relate this to the Lebesgue integral and show how it extends to our more general context.

Corollary 3.8 *Suppose $f(x)$ is a non-negative function on $\mathbb{R}^d$, and let*

$$\mathcal{A} = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} : \ 0 \leq y \leq f(x)\}.$$

Then:

(i) *f is measurable on $\mathbb{R}^d$ if and only if $\mathcal{A}$ is measurable in $\mathbb{R}^{d+1}$.*

(ii) *If the conditions in* (i) *hold, then*

$$\int_{\mathbb{R}^d} f(x)\, dx = m(\mathcal{A}).$$

Proof. If f is measurable on $\mathbb{R}^d$, then the previous proposition guarantees that the function

$$F(x, y) = y - f(x)$$

is measurable on $\mathbb{R}^{d+1}$, so we immediately see that $\mathcal{A} = \{y \geq 0\} \cap \{F \leq 0\}$ is measurable.

Conversely, suppose that $\mathcal{A}$ is measurable. We note that for each $x \in \mathbb{R}^{d_1}$ the slice $\mathcal{A}_x = \{y \in \mathbb{R} : (x, y) \in \mathcal{A}\}$ is a closed segment, namely $\mathcal{A}_x = [0, f(x)]$. Consequently Corollary 3.3 (with the roles of x and y interchanged) yields the measurability of $m(\mathcal{A}_x) = f(x)$. Moreover

$$m(\mathcal{A}) = \int \chi_{\mathcal{A}}(x, y)\, dx\, dy = \int_{\mathbb{R}^{d_1}} m(\mathcal{A}_x)\, dx = \int_{\mathbb{R}^{d_1}} f(x)\, dx,$$

as was to be shown.

We conclude this section with a useful result.

Proposition 3.9 *If f is a measurable function on $\mathbb{R}^d$, then the function $\tilde{f}(x, y) = f(x - y)$ is measurable on $\mathbb{R}^d \times \mathbb{R}^d$.*

By picking $E = \{z \in \mathbb{R}^d : f(z) < a\}$, we see that it suffices to prove that whenever E is a measurable subset of $\mathbb{R}^d$, then $\tilde{E} = \{(x, y) : x - y \in E\}$ is a measurable subset of $\mathbb{R}^d \times \mathbb{R}^d$.

Note first that if $\mathcal{O}$ is an open set, then $\tilde{\mathcal{O}}$ is also open. Taking countable intersections shows that if E is a G_δ set, then so is $\tilde{E}$. Assume now that $m(\tilde{E}_k) = 0$ for each k, where $\tilde{E}_k = \tilde{E} \cap B_k$ and $B_k = \{|y| < k\}$. Again, take $\mathcal{O}$ to be open in $\mathbb{R}^d$, and let us calculate $m(\tilde{\mathcal{O}} \cap B_k)$. We have that $\chi_{\tilde{\mathcal{O}} \cap B_k} = \chi_{\mathcal{O}}(x - y)\chi_{B_k}(y)$. Hence

$$\begin{aligned} m(\tilde{\mathcal{O}} \cap B_k) &= \int \chi_{\mathcal{O}}(x - y)\chi_{B_k}(y)\, dy\, dx \\ &= \int \left(\int \chi_{\mathcal{O}}(x - y)\, dx \right) \chi_{B_k}(y)\, dy \\ &= m(\mathcal{O})\, m(B_k), \end{aligned}$$

by the translation-invariance of the measure. Now if $m(E) = 0$, there is a sequence of open sets $\mathcal{O}_n$ such that $E \subset \mathcal{O}_n$ and $m(\mathcal{O}_n) \to 0$. It follows from the above that $\tilde{E}_k \subset \tilde{\mathcal{O}}_n \cap B_k$ and $m(\tilde{\mathcal{O}}_n \cap B_k) \to 0$ in n for each fixed k. This shows $m(\tilde{E}_k) = 0$, and hence $m(\tilde{E}) = 0$. The proof of the proposition is concluded once we recall that any measurable set E can be written as the difference of a G_δ and a set of measure zero.

4* A Fourier inversion formula

The question of the inversion of the Fourier transform encompasses in effect the problem at the origin of Fourier analysis. This issue involves

establishing the validity of the inversion formula for a function f in terms of its Fourier transform $\hat{f}$, that is,

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \cdot \xi}\, dx, \tag{14}$$

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i x \cdot \xi}\, d\xi. \tag{15}$$

We have already encountered this problem in Book I in the rudimentary case when in fact both f and $\hat{f}$ were continuous and had rapid (or moderate) decrease at infinity. In Book II we also considered the question in the one-dimensional setting, seen from the viewpoint of complex analysis. The most elegant and useful formulations of Fourier inversion are in terms of the L^2 theory, or in its greatest generality stated in the language of distributions. We shall take up these matters systematically later.[4] It will, nevertheless, be enlightening to digress here to see what our knowledge at this stage teaches us about this problem. We intend to do this by presenting a variant of the inversion formula appropriate for L^1, one that is both simple and adequate in many circumstances.

To begin with, we need to have an idea of what can be said about the Fourier transform of an arbitrary function in $L^1(\mathbb{R}^d)$.

Proposition 4.1 *Suppose $f \in L^1(\mathbb{R}^d)$. Then $\hat{f}$ defined by* (14) *is continuous and bounded on $\mathbb{R}^d$.*

In fact, since $|f(x)e^{-2\pi i x \cdot \xi}| = |f(x)|$, the integral representing $\hat{f}$ converges for each ξ and $\sup_{\xi \in \mathbb{R}^d} |\hat{f}(\xi)| \leq \int_{\mathbb{R}^d} |f(x)|\, dx = \|f\|$. To verify the continuity, note that for every x, $f(x)e^{-2\pi i x \cdot \xi} \to f(x)e^{-2\pi i x \cdot \xi_0}$ as $\xi \to \xi_0$, where ξ_0 is any point in $\mathbb{R}^d$; hence $\hat{f}(\xi) \to \hat{f}(\xi_0)$ by the dominated convergence theorem.

One can assert a little more than the boundedness of $\hat{f}$; namely, one has $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$, but not much more can be said about the decrease at infinity of $\hat{f}$. (See Exercises 22 and 25.) As a consequence, for general $f \in L^1(\mathbb{R}^d)$ the function $\hat{f}$ is not in $L^1(\mathbb{R}^d)$, and the presumed formula (15) becomes problematical. The following theorem evades this difficulty and yet is useful in a number of situations.

Theorem 4.2 *Suppose $f \in L^1(\mathbb{R}^d)$ and assume also that $\hat{f} \in L^1(\mathbb{R}^d)$. Then the inversion formula* (15) *holds for almost every x.*

An immediate corollary is the uniqueness of the Fourier transform on L^1.

[4] The L^2 theory will be dealt with in Chapter 5, and distributions will be studied in Book IV.

Corollary 4.3 *Suppose $\hat{f}(\xi) = 0$ for all ξ. Then $f = 0$ a.e.*

The proof of the theorem requires only that we adapt the earlier arguments carried out for Schwartz functions in Chapter 5 of Book I to the present context. We begin with the "multiplication formula."

Lemma 4.4 *Suppose f and g belong to $L^1(\mathbb{R}^d)$. Then*

$$\int_{\mathbb{R}^d} \hat{f}(\xi)g(\xi)\, d\xi = \int_{\mathbb{R}^d} f(y)\hat{g}(y)\, dy.$$

Note that both integrals converge in view of the proposition above. Consider the function $F(\xi, y) = g(\xi)f(y)e^{-2\pi i\xi\cdot y}$ defined for $(\xi, y) \in \mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{2d}$. It is measurable as a function on $\mathbb{R}^{2d}$ in view of Corollary 3.7. We now apply Fubini's theorem to observe first that

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |F(\xi, y)|\, d\xi\, dy = \int_{\mathbb{R}^d} |g(\xi)|\, d\xi \int_{\mathbb{R}^d} |f(y)|\, dy < \infty.$$

Next, if we evaluate $\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} F(\xi, y)\, d\xi\, dy$ by writing it as $\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} F(\xi, y)\, d\xi\right) dy$ we get the left-hand side of the desired equality. Evaluating the double integral in the reverse order gives as the right-hand side, proving the lemma.

Next we consider the modulated Gaussian, $g(\xi) = e^{-\pi\delta|\xi|^2}e^{2\pi i x\cdot\xi}$, where for the moment δ and x are fixed, with $\delta > 0$ and $x \in \mathbb{R}^d$. An elementary calculation gives[5]

$$\hat{g}(y) = \int_{\mathbb{R}^d} e^{-\pi\delta|\xi|^2}e^{2\pi i(x-y)\cdot\xi}\, d\xi = \delta^{-d/2}e^{-\pi|x-y|^2/\delta},$$

which we will abbreviate as $K_\delta(x - y)$. We recognize K_δ as a "good kernel" that satisfies:

(i) $\displaystyle\int_{\mathbb{R}^d} K_\delta(y)\, dy = 1.$

(ii) For each $\eta > 0$, $\displaystyle\int_{|y|>\eta} K_\delta(y)\, dy \to 0$ as $\delta \to 0$.

Applying the lemma gives

$$\int_{\mathbb{R}^d} \hat{f}(\xi)e^{-\pi\delta|\xi|^2}e^{2\pi i x\cdot\xi}\, d\xi = \int_{\mathbb{R}^d} f(y)K_\delta(x - y)\, dy. \tag{16}$$

[5]See for example Chapter 6 in Book I.

Note that since $\hat{f} \in L^1(\mathbb{R}^d)$, the dominated convergence theorem shows that the left-hand side of (16) converges to $\int_{\mathbb{R}^d} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\, d\xi$ as $\delta \to 0$, for each x. As for the right-hand side, we make two successive change of variables $y \to y + x$ (a translation), and $y \to -y$ (a reflection), and take into account the corresponding invariance of the integrals (see equations (4) and (5)). Thus the right-hand side becomes $\int_{\mathbb{R}^d} f(x-y)K_\delta(y)\, dy$, and we will prove that this function converges in the L^1-norm to f as $\delta \to 0$. Indeed, we can write the difference as

$$\Delta_\delta(x) = \int_{\mathbb{R}^d} f(x-y)K_\delta(y)\, dy - f(x) = \int_{\mathbb{R}^d} (f(x-y) - f(x))K_\delta(y)\, dy,$$

because of property (i) above. Thus

$$|\Delta_\delta(x)| \leq \int_{\mathbb{R}^d} |f(x-y) - f(x)|K_\delta(y)\, dy.$$

We can now apply Fubini's theorem, recalling that the measurability of $f(x)$ and $f(x-y)$ on $\mathbb{R}^d \times \mathbb{R}^d$ are established in Corollary 3.7 and Proposition 3.9. The result is

$$\|\Delta_\delta\| \leq \int_{\mathbb{R}^d} \|f_y - f\|K_\delta(y)\, dy, \qquad \text{where } f_y(x) = f(x-y).$$

Now, for given $\epsilon > 0$ we can find (by Proposition 2.5) $\eta > 0$ so small such that $\|f_y - f\| < \epsilon$ when $|y| < \eta$. Thus

$$\|\Delta_\delta\| \leq \epsilon + \int_{|y|>\eta} \|f_y - f\|K_\delta(y)\, dy \leq \epsilon + 2\|f\| \int_{|y|>\eta} K_\delta(y)\, dy.$$

The first inequality follows by using (i) again; the second holds because $\|f_y - f\| \leq \|f_y\| + \|f\| = 2\|f\|$. Therefore, with the use of (ii), the combination above is $\leq 2\epsilon$ if δ is sufficiently small. To summarize: the right-hand side of (16) converges to f in the L^1-norm as $\delta \to 0$, and thus by Corollary 2.3 there is a subsequence that converges to $f(x)$ almost everywhere, and the theorem is proved.

Note that an immediate consequence of the theorem and the proposition is that if $\hat{f}$ were in L^1, then f could be modified on a set of measure zero to become continuous everywhere. This is of course impossible for the general $f \in L^1(\mathbb{R}^d)$.

5 Exercises

1. Given a collection of sets $F_1, F_2, \ldots, F_n$, construct another collection $F_1^*, F_2^*, \ldots, F_N^*$, with $N = 2^n - 1$, so that $\bigcup_{k=1}^n F_k = \bigcup_{j=1}^N F_j^*$; the collection $\{F_j^*\}$ is disjoint; also

$F_k = \bigcup_{F_j^* \subset F_k} F_j^*$, for every k.

[Hint: Consider the 2^n sets $F_1' \cap F_2' \cap \cdots \cap F_n'$ where each F_k' is either F_k or F_k^c.]

2. In analogy to Proposition 2.5, prove that if f is integrable on $\mathbb{R}^d$ and $\delta > 0$, then $f(\delta x)$ converges to $f(x)$ in the L^1-norm as $\delta \to 1$.

3. Suppose f is integrable on $(-\pi, \pi]$ and extended to $\mathbb{R}$ by making it periodic of period 2π. Show that

$$\int_{-\pi}^{\pi} f(x)\,dx = \int_I f(x)\,dx,$$

where I is any interval in $\mathbb{R}$ of length 2π.

[Hint: I is contained in two consecutive intervals of the form $(k\pi, (k+2)\pi)$.]

4. Suppose f is integrable on $[0, b]$, and

$$g(x) = \int_x^b \frac{f(t)}{t}\,dt \qquad \text{for } 0 < x \leq b.$$

Prove that g is integrable on $[0, b]$ and

$$\int_0^b g(x)\,dx = \int_0^b f(t)\,dt.$$

5. Suppose F is a closed set in $\mathbb{R}$, whose complement has finite measure, and let $\delta(x)$ denote the distance from x to F, that is,

$$\delta(x) = d(x, F) = \inf\{|x - y| : \ y \in F\}.$$

Consider

$$I(x) = \int_{\mathbb{R}} \frac{\delta(y)}{|x - y|^2}\,dy.$$

(a) Prove that δ is continuous, by showing that it satisfies the Lipschitz condition

$$|\delta(x) - \delta(y)| \leq |x - y|.$$

(b) Show that $I(x) = \infty$ for each $x \notin F$.

(c) Show that $I(x) < \infty$ for a.e. $x \in F$. This may be surprising in view of the fact that the Lispshitz condition cancels only one power of $|x - y|$ in the integrand of I.

[Hint: For the last part, investigate $\int_F I(x)\,dx$.]

6. Integrability of f on $\mathbb{R}$ does not necessarily imply the convergence of $f(x)$ to 0 as $x \to \infty$.

(a) There exists a positive continuous function f on $\mathbb{R}$ so that f is integrable on $\mathbb{R}$, but yet $\limsup_{x\to\infty} f(x) = \infty$.

(b) However, if we assume that f is uniformly continuous on $\mathbb{R}$ and integrable, then $\lim_{|x|\to\infty} f(x) = 0$.

[Hint: For (a), construct a continuous version of the function equal to n on the segment $[n, n+1/n^3)$, $n \geq 1$.]

7. Let $\Gamma \subset \mathbb{R}^d \times \mathbb{R}$, $\Gamma = \{(x,y) \in \mathbb{R}^d \times \mathbb{R} : \ y = f(x)\}$, and assume f is measurable on $\mathbb{R}^d$. Show that Γ is a measurable subset of $\mathbb{R}^{d+1}$, and $m(\Gamma) = 0$.

8. If f is integrable on $\mathbb{R}$, show that $F(x) = \int_{-\infty}^{x} f(t)\,dt$ is uniformly continuous.

9. Tchebychev inequality. Suppose $f \geq 0$, and f is integrable. If $\alpha > 0$ and $E_\alpha = \{x : f(x) > \alpha\}$, prove that

$$m(E_\alpha) \leq \frac{1}{\alpha} \int f.$$

10. Suppose $f \geq 0$, and let $E_{2^k} = \{x : f(x) > 2^k\}$ and $F_k = \{x : 2^k < f(x) \leq 2^{k+1}\}$. If f is finite almost everywhere, then

$$\bigcup_{k=-\infty}^{\infty} F_k = \{f(x) > 0\},$$

and the sets F_k are disjoint.

Prove that f is integrable if and only if

$$\sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty, \qquad \text{if and only if} \qquad \sum_{k=-\infty}^{\infty} 2^k m(E_{2^k}) < \infty.$$

Use this result to verify the following assertions. Let

$$f(x) = \begin{cases} |x|^{-a} & \text{if } |x| \leq 1, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad g(x) = \begin{cases} |x|^{-b} & \text{if } |x| > 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then f is integrable on $\mathbb{R}^d$ if and only if $a < d$; also g is integrable on $\mathbb{R}^d$ if and only if $b > d$.

11. Prove that if f is integrable on $\mathbb{R}^d$, real-valued, and $\int_E f(x)\,dx \geq 0$ for every measurable E, then $f(x) \geq 0$ a.e. x. As a result, if $\int_E f(x)\,dx = 0$ for every measurable E, then $f(x) = 0$ a.e.

12. Show that there are $f \in L^1(\mathbb{R}^d)$ and a sequence $\{f_n\}$ with $f_n \in L^1(\mathbb{R}^d)$ such that

$$\|f - f_n\|_{L^1} \to 0,$$

but $f_n(x) \to f(x)$ for no x.

[Hint: In $\mathbb{R}$, let $f_n = \chi_{I_n}$, where I_n is an appropriately chosen sequence of intervals with $m(I_n) \to 0$.]

13. Give an example of two measurable sets A and B such that $A + B$ is not measurable.

[Hint: In $\mathbb{R}^2$ take $A = \{0\} \times [0, 1]$ and $B = \mathcal{N} \times \{0\}$.]

14. In Exercise 6 of the previous chapter we saw that $m(B) = v_d r^d$, whenever B is a ball of radius r in $\mathbb{R}^d$ and $v_d = m(B_1)$, with B_1 the unit ball. Here we evaluate the constant v_d.

(a) For $d = 2$, prove using Corollary 3.8 that

$$v_2 = 2\int_{-1}^{1} (1 - x^2)^{1/2}\, dx,$$

and hence by elementary calculus, that $v_2 = \pi$.

(b) By similar methods, show that

$$v_d = 2v_{d-1}\int_0^1 (1 - x^2)^{(d-1)/2}\, dx.$$

(c) The result is

$$v_d = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)}.$$

Another derivation is in Exercise 5 in Chapter 6 below. Relevant facts about the gamma and beta functions can be found in Chapter 6 of Book II.

15. Consider the function defined over $\mathbb{R}$ by

$$f(x) = \begin{cases} x^{-1/2} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

For a fixed enumeration $\{r_n\}_{n=1}^{\infty}$ of the rationals $\mathbb{Q}$, let

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} f(x - r_n).$$

Prove that F is integrable, hence the series defining F converges for almost every $x \in \mathbb{R}$. However, observe that this series is unbounded on every interval, and in fact, any function $\tilde{F}$ that agrees with F a.e is unbounded in any interval.

16. Suppose f is integrable on $\mathbb{R}^d$. If $\delta = (\delta_1, \ldots, \delta_d)$ is a d-tuple of non-zero real numbers, and

$$f^\delta(x) = f(\delta x) = f(\delta_1 x_1, \ldots, \delta_d x_d),$$

show that f^δ is integrable with

$$\int_{\mathbb{R}^d} f^\delta(x)\, dx = |\delta_1|^{-1} \cdots |\delta_d|^{-1} \int_{\mathbb{R}^d} f(x)\, dx.$$

17. Suppose f is defined on $\mathbb{R}^2$ as follows: $f(x,y) = a_n$ if $n \le x < n+1$ and $n \le y < n+1$, $(n \ge 0)$; $f(x,y) = -a_n$ if $n \le x < n+1$ and $n+1 \le y < n+2$, $(n \ge 0)$; while $f(x,y) = 0$ elsewhere. Here $a_n = \sum_{k \le n} b_k$, with $\{b_k\}$ a positive sequence such that $\sum_{k=0}^{\infty} b_k = s < \infty$.

(a) Verify that each slice f^y and f_x is integrable. Also for all x, $\int f_x(y)\, dy = 0$, and hence $\int \left(\int f(x,y)\, dy \right) dx = 0$.

(b) However, $\int f^y(x)\, dx = a_0$ if $0 \le y < 1$, and $\int f^y(x)\, dx = a_n - a_{n-1}$ if $n \le y < n+1$ with $n \ge 1$. Hence $y \mapsto \int f^y(x)\, dx$ is integrable on $(0, \infty)$ and

$$\int \left(\int f(x,y)\, dx \right) dy = s.$$

(c) Note that $\int_{\mathbb{R}\times\mathbb{R}} |f(x,y)|\, dx\, dy = \infty$.

18. Let f be a measurable finite-valued function on $[0,1]$, and suppose that $|f(x) - f(y)|$ is integrable on $[0,1] \times [0,1]$. Show that $f(x)$ is integrable on $[0,1]$.

19. Suppose f is integrable on $\mathbb{R}^d$. For each $\alpha > 0$, let $E_\alpha = \{x : |f(x)| > \alpha\}$. Prove that

$$\int_{\mathbb{R}^d} |f(x)|\, dx = \int_0^\infty m(E_\alpha)\, d\alpha.$$

20. The problem (highlighted in the discussion preceding Fubini's theorem) that certain slices of measurable sets can be non-measurable may be avoided by restricting attention to Borel measurable functions and Borel sets. In fact, prove the following:

Suppose E is a Borel set in $\mathbb{R}^2$. Then for every y, the slice E^y is a Borel set in $\mathbb{R}$.

[Hint: Consider the collection $\mathcal{C}$ of subsets E of $\mathbb{R}^2$ with the property that each slice E^y is a Borel set in $\mathbb{R}$. Verify that $\mathcal{C}$ is a σ-algebra that contains the open sets.]

21. Suppose that f and g are measurable functions on $\mathbb{R}^d$.

(a) Prove that $f(x-y)g(y)$ is measurable on $\mathbb{R}^{2d}$.

(b) Show that if f and g are integrable on $\mathbb{R}^d$, then $f(x-y)g(y)$ is integrable on $\mathbb{R}^{2d}$.

(c) Recall the definition of the convolution of f and g given by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\, dy.$$

Show that $f * g$ is well defined for a.e. x (that is, $f(x-y)g(y)$ is integrable on $\mathbb{R}^d$ for a.e. x).

(d) Show that $f * g$ is integrable whenever f and g are integrable, and that

$$\|f * g\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}\, \|g\|_{L^1(\mathbb{R}^d)},$$

with equality if f and g are non-negative.

(e) The Fourier transform of an integrable function f is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}\, dx.$$

Check that $\hat{f}$ is bounded and is a continuous function of ξ. Prove that for each ξ one has

$$\widehat{(f * g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

22. Prove that if $f \in L^1(\mathbb{R}^d)$ and

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\xi}\, dx,$$

then $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$. (This is the Riemann-Lebesgue lemma.)
[Hint: Write $\hat{f}(\xi) = \frac{1}{2}\int_{\mathbb{R}^d}[f(x) - f(x-\xi')]e^{-2\pi i x\xi}\, dx$, where $\xi' = \frac{1}{2}\frac{\xi}{|\xi|^2}$, and use Proposition 2.5.]

23. As an application of the Fourier transform, show that there does not exist a function $I \in L^1(\mathbb{R}^d)$ such that

$$f * I = f \quad \text{for all } f \in L^1(\mathbb{R}^d).$$

24. Consider the convolution

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\, dy.$$

(a) Show that $f * g$ is uniformly continuous when f is integrable and g bounded.

(b) If in addition g is integrable, prove that $(f * g)(x) \to 0$ as $|x| \to \infty$.

25. Show that for each $\epsilon > 0$ the function $F(\xi) = \frac{1}{(1+|\xi|^2)^\epsilon}$ is the Fourier transform of an L^1 function.

[Hint: With $K_\delta(x) = e^{-\pi|x|^2/\delta}\delta^{-d/2}$ consider $f(x) = \int_0^\infty K_\delta(x)e^{-\pi\delta}\delta^{\epsilon-1}\, d\delta$. Use Fubini's theorem to prove $f \in L^1(\mathbb{R}^d)$, and

$$\hat{f}(\xi) = \int_0^\infty e^{-\pi\delta|\xi|^2} e^{-\pi\delta}\delta^{\epsilon-1}\, d\delta,$$

and evaluate the last integral as $\pi^{-\epsilon}\Gamma(\epsilon)\frac{1}{(1+|\xi|^2)^\epsilon}$. Here $\Gamma(s)$ is the gamma function defined by $\Gamma(s) = \int_0^\infty e^{-t}t^{s-1}\, dt$.]

6 Problems

1. If f is integrable on $[0, 2\pi]$, then $\int_0^{2\pi} f(x)e^{-inx}\, dx \to 0$ as $|n| \to \infty$.
Show as a consequence that if E is a measurable subset of $[0, 2\pi]$, then

$$\int_E \cos^2(nx + u_n)\, dx \to \frac{m(E)}{2}, \qquad \text{as } n \to \infty$$

for any sequence $\{u_n\}$.

[Hint: See Exercise 22.]

2. Prove the Cantor-Lebesgue theorem: if

$$\sum_{n=0}^{\infty} A_n(x) = \sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx)$$

converges for x in a set of positive measure (or in particular for all x), then $a_n \to 0$ and $b_n \to 0$ as $n \to \infty$.

[Hint: Note that $A_n(x) \to 0$ uniformly on a set E of positive measure.]

3. A sequence $\{f_k\}$ of measurable functions on $\mathbb{R}^d$ is **Cauchy in measure** if for every $\epsilon > 0$,

$$m(\{x : |f_k(x) - f_\ell(x)| > \epsilon\}) \to 0 \qquad \text{as } k, \ell \to \infty.$$

We say that $\{f_k\}$ **converges in measure** to a (measurable) function f if for every $\epsilon > 0$

$$m(\{x : |f_k(x) - f(x)| > \epsilon\}) \to 0 \qquad \text{as } k \to \infty.$$

This notion coincides with the "convergence in probability" of probability theory.

Prove that if a sequence $\{f_k\}$ of integrable functions converges to f in L^1, then $\{f_k\}$ converges to f in measure. Is the converse true?

We remark that this mode of convergence appears naturally in the proof of Egorov's theorem.

4. We have already seen (in Exercise 8, Chapter 1) that if E is a measurable set in $\mathbb{R}^d$, and L is a linear transformation of $\mathbb{R}^d$ to $\mathbb{R}^d$, then $L(E)$ is also measurable, and if E has measure 0, then so has $L(E)$. The quantitative statement is

$$m(L(E)) = |\det(L)|\, m(E).$$

As a special case, note that the Lebesgue measure is invariant under rotations. (For this special case see also Exercise 26 in the next chapter.)

The above identity can be proved using Fubini's theorem as follows.

(a) Consider first the case $d = 2$, and L a "strictly" upper triangular transformation $x' = x + ay$, $y' = y$. Then

$$\chi_{L(E)}(x, y) = \chi_E(L^{-1}(x, y)) = \chi_E(x - ay, y).$$

Hence

$$\begin{aligned} m(L(E)) &= \int_{\mathbb{R}\times\mathbb{R}} \left(\int \chi_E(x - ay, y) \right) dy \\ &= \int_{\mathbb{R}\times\mathbb{R}} \left(\int \chi_E(x, y)\, dx \right) dy \\ &= m(E), \end{aligned}$$

by the translation-invariance of the measure.

(b) Similarly $m(L(E)) = m(E)$ if L is strictly lower triangular. In general, one can write $L = L_1 \Delta L_2$, where L_j are strictly (upper and lower) triangular and Δ is diagonal. Thus $m(L(E)) = |\det(L)| m(E)$, if one uses Exercise 7 in Chapter 1.

5. There is an ordering $\prec$ of $\mathbb{R}$ with the property that for each $y \in \mathbb{R}$ the set $\{x \in \mathbb{R} :\ x \prec y\}$ is at most countable.

The existence of this ordering depends on the **continuum hypothesis**, which asserts: whenever S is an infinite subset of $\mathbb{R}$, then either S is countable, or S has the cardinality of $\mathbb{R}$ (that is, can be mapped bijectively to $\mathbb{R}$).[6]

[6] This assertion, formulated by Cantor, is like the well-ordering principle independent of the other axioms of set theory, and so we are also free to accept its validity.

[Hint: Let $\prec$ denote a well-ordering of $\mathbb{R}$, and define the set X by $X = \{y \in \mathbb{R} : \text{the set } \{x : x \prec y\} \text{ is not countable}\}$. If X is empty we are done. Otherwise, consider the smallest element $\overline{y}$ in X, and use the continuum hypothesis.]

3 Differentiation and Integration

> The Maximal Problem:
> The problem is most easily grasped when stated in the language of cricket, or any other game in which a player compiles a series of scores of which an average is recorded.
>
> *G. H. Hardy and J. E. Littlewood,* 1930

That differentiation and integration are inverse operations was already understood early in the study of the calculus. Here we want to reexamine this basic idea in the framework of the general theory studied in the previous chapters. Our objective is the formulation and proof of the fundamental theorem of the calculus in this setting, and the development of some of the concepts that occur. We shall try to achieve this by answering two questions, each expressing one of the ways of representing the reciprocity between differentiation and integration.

The first problem involved may be stated as follows.

- Suppose f is integrable on $[a,b]$ and F is its indefinite integral $F(x) = \int_a^x f(y)\,dy$. Does this imply that F is differentiable (at least for almost every x), and that $F' = f$?

We shall see that the affirmative answer to this question depends on ideas that have broad application and are not limited to the one-dimensional situation.

For the second question we reverse the order of differentiation and integration.

- What conditions on a function F on $[a,b]$ guarantee that $F'(x)$ exists (for a.e. x), that this function is integrable, and that moreover

$$F(b) - F(a) = \int_a^b F'(x)\,dx \quad ?$$

While this problem will be examined from a narrower perspective than the first, the issues it raises are deep and the consequences entailed are

far-reaching. In particular, we shall find that this question is connected to the problem of rectifiability of curves, and as an illustration of this link, we shall establish the general isoperimetric inequality in the plane.

1 Differentiation of the integral

We begin with the first problem, that is, the study of differentiation of the integral. If f is given on $[a, b]$ and integrable on that interval, we let

$$F(x) = \int_a^x f(y)\, dy, \qquad a \leq x \leq b.$$

To deal with $F'(x)$, we recall the definition of the derivative as the limit of the quotient

$$\frac{F(x+h) - F(x)}{h} \qquad \text{when } h \text{ tends to } 0.$$

We note that this quotient takes the form (say in the case $h > 0$)

$$\frac{1}{h}\int_x^{x+h} f(y)\, dy = \frac{1}{|I|}\int_I f(y)\, dy,$$

where we use the notation $I = (x, x+h)$ and $|I|$ for the length of this interval. At this point, we pause to observe that the above expression is the "average" value of f over I, and that in the limit as $|I| \to 0$, we might expect that these averages tend to $f(x)$. Reformulating the question slightly, we may ask whether

$$\lim_{\substack{|I| \to 0 \\ x \in I}} \frac{1}{|I|}\int_I f(y)\, dy = f(x)$$

holds for suitable points x. In higher dimensions we can pose a similar question, where the averages of f are taken over appropriate sets that generalize the intervals in one dimension. Initially we shall study this problem where the sets involved are the balls B containing x, with their volume $m(B)$ replacing the length $|I|$ of I. Later we shall see that as a consequence of this special case similar results will hold for more general collections of sets, those that have bounded "eccentricity."

With this in mind we restate our first problem in the context of $\mathbb{R}^d$, for all $d \geq 1$.

Suppose f is integrable on $\mathbb{R}^d$. Is it true that

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\,dy = f(x), \qquad \text{for a.e. } x?$$

The limit is taken as the volume of open balls B containing x goes to 0.

We shall refer to this question as the **averaging problem**. We remark that if B is any ball of radius r in $\mathbb{R}^d$, then $m(B) = v_d r^d$, where v_d is the measure of the unit ball. (See Exercise 14 in the previous chapter.)

Note of course that in the special case when f is continuous at x , the limit does converge to $f(x)$. Indeed, given $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$. Since

$$f(x) - \frac{1}{m(B)} \int_B f(y)\,dy = \frac{1}{m(B)} \int_B (f(x) - f(y))\,dy,$$

we find that whenever B is a ball of radius $< \delta/2$ that contains x, then

$$\left| f(x) - \frac{1}{m(B)} \int_B f(y)\,dy \right| \leq \frac{1}{m(B)} \int_B |f(x) - f(y)|\,dy < \epsilon,$$

as desired.

The averaging problem has an affirmative answer, but to establish that fact, which is qualitative in nature, we need to make some quantitative estimates bearing on the overall behavior of the averages of f. This will be done in terms of the maximal averages of $|f|$, to which we now turn.

1.1 The Hardy-Littlewood maximal function

The maximal function that we consider below arose first in the one-dimensional situation treated by Hardy and Littlewood. It seems that they were led to the study of this function by toying with the question of how a batsman's score in cricket may best be distributed to maximize his satisfaction. As it turns out, the concepts involved have a universal significance in analysis. The relevant definition is as follows.

If f is integrable on $\mathbb{R}^d$, we define its **maximal function** f^* by

$$f^*(x) = \sup_{x \in B} \frac{1}{m(B)} \int_B |f(y)|\,dy, \qquad x \in \mathbb{R}^d,$$

where the supremum is taken over all balls containing the point x. In other words, we replace the limit in the statement of the averaging problem by a supremum, and f by its absolute value.

The main properties of f^* we shall need are summarized in a theorem.

Theorem 1.1 *Suppose f is integrable on $\mathbb{R}^d$. Then:*

(i) *f^* is measurable.*

(ii) *$f^*(x) < \infty$ for a.e. x.*

(iii) *f^* satisfies*

$$m(\{x \in \mathbb{R}^d : \; f^*(x) > \alpha\}) \leq \frac{A}{\alpha} \, \|f\|_{L^1(\mathbb{R}^d)} \tag{1}$$

for all $\alpha > 0$, where $A = 3^d$, and $\|f\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f(x)| \, dx$.

Before we come to the proof we want to clarify the nature of the main conclusion (iii). As we shall observe, one has that $f^*(x) \geq |f(x)|$ for a.e. x; the effect of (iii) is that, broadly speaking, f^* is not much larger than $|f|$. From this point of view, we would have liked to conclude that f^* is integrable, as a result of the assumed integrability of f. However, this is not the case, and (iii) is the best substitute available (see Exercises 4 and 5).

An inequality of the type (1) is called a **weak-type** inequality because it is weaker than the corresponding inequality for the L^1-norms. Indeed, this can be seen from the Tchebychev inequality (Exercise 9 in Chapter 2), which states that for an arbitrary integrable function g,

$$m(\{x : \; |g(x)| > \alpha\}) \leq \frac{1}{\alpha} \, \|g\|_{L^1(\mathbb{R}^d)}, \qquad \text{for all } \alpha > 0.$$

We should add that the exact value of A in the inequality (1) is unimportant for us. What matters is that this constant be independent of α and f.

The only simple assertion in the theorem is that f^* is a measurable function. Indeed, the set $E_\alpha = \{x \in \mathbb{R}^d : \; f^*(x) > \alpha\}$ is open, because if $\overline{x} \in E_\alpha$, there exists a ball B such that $\overline{x} \in B$ and

$$\frac{1}{m(B)} \int_B |f(y)| \, dy > \alpha.$$

Now any point x close enough to $\overline{x}$ will also belong to B; hence $x \in E_\alpha$ as well.

The two other properties of f^* in the theorem are deeper, with (ii) being a consequence of (iii). This follows at once if we observe that

$$\{x : \; f^*(x) = \infty\} \subset \{x : \; f^*(x) > \alpha\}$$

for all α. Taking the limit as α tends to infinity, the third property yields $m(\{x: \ f^*(x) = \infty\}) = 0$.

The proof of inequality (1) relies on an elementary version of a Vitali covering argument.[1]

Lemma 1.2 *Suppose $\mathcal{B} = \{B_1, B_2, \ldots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^d$. Then there exists a disjoint sub-collection $B_{i_1}, B_{i_2}, \ldots, B_{i_k}$ of $\mathcal{B}$ that satisfies*

$$m\left(\bigcup_{\ell=1}^{N} B_\ell\right) \leq 3^d \sum_{j=1}^{k} m(B_{i_j}).$$

Loosely speaking, we may always find a disjoint sub-collection of balls that covers a fraction of the region covered by the original collection of balls.

Proof. The argument we give is constructive and relies on the following simple observation: Suppose B and B' are a pair of balls that intersect, with the radius of B' being not greater than that of B. Then B' is contained in the ball $\tilde{B}$ that is concentric with B but with 3 times its radius.

As a first step, we pick a ball B_{i_1} in $\mathcal{B}$ with maximal (that is, largest) radius, and then delete from $\mathcal{B}$ the ball B_{i_1} as well as any balls that intersect B_{i_1}. Thus all the balls that are deleted are contained in the ball $\tilde{B}_{i_1}$ concentric with B_{i_1}, but with 3 times its radius.

The remaining balls yield a new collection $\mathcal{B}'$, for which we repeat the procedure. We pick B_{i_2} with largest radius in $\mathcal{B}'$, and then delete from $\mathcal{B}'$ the ball B_{i_2} and any ball that intersects B_{i_2}. Continuing this way we find, after at most N steps, a collection of disjoint balls $B_{i_1}, B_{i_2}, \ldots, B_{i_k}$.

Finally, to prove that this disjoint collection of balls satisfies the inequality in the lemma, we use the observation made at the beginning of the proof. We let $\tilde{B}_{i_j}$ denote the ball concentric with B_{i_j}, but with 3 times its radius. Since any ball B in $\mathcal{B}$ must intersect a ball B_{i_j} and have equal or smaller radius than B_{i_j}, we must have $B \subset \tilde{B}_{i_j}$, thus

$$m\left(\bigcup_{\ell=1}^{N} B_\ell\right) \leq m\left(\bigcup_{j=1}^{k} \tilde{B}_{i_j}\right) \leq \sum_{j=1}^{k} m(\tilde{B}_{i_j}) = 3^d \sum_{j=1}^{k} m(B_{i_j}).$$

---

[1] We note that the lemma that follows is the first of a series of covering arguments that occur below in the theory of differentiation; see also Lemma 3.9 and its corollary, as well as Lemma 3.5, where the covering assertion is more implicit.

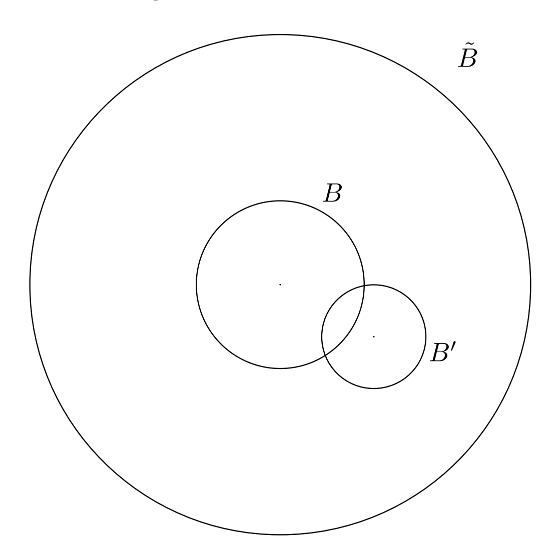


Figure 1. The balls B and $\tilde{B}$

In the last step we have used the fact that in $\mathbb{R}^d$ a dilation of a set by $\delta > 0$ results in the multiplication by δ^d of the Lebesgue measure of this set.

The proof of (iii) in Theorem 1.1 is now in reach. If we let $E_\alpha = \{x : f^*(x) > \alpha\}$, then for each $x \in E_\alpha$ there exists a ball B_x that contains x, and such that

$$\frac{1}{m(B_x)} \int_{B_x} |f(y)|\, dy > \alpha.$$

Therefore, for each ball B_x we have

$$m(B_x) < \frac{1}{\alpha} \int_{B_x} |f(y)|\, dy. \tag{2}$$

Fix a compact subset K of E_α. Since K is covered by $\bigcup_{x \in E_\alpha} B_x$, we may select a finite subcover of K, say $K \subset \bigcup_{\ell=1}^N B_\ell$. The covering lemma guarantees the existence of a sub-collection $B_{i_1}, \ldots, B_{i_k}$ of disjoint balls with

$$m\left(\bigcup_{\ell=1}^N B_\ell\right) \leq 3^d \sum_{j=1}^k m(B_{i_j}). \tag{3}$$

Since the balls $B_{i_1}, \ldots, B_{i_k}$ are disjoint and satisfy (2) as well as (3), we find that

$$\begin{aligned} m(K) \leq m\left(\bigcup_{\ell=1}^{N} B_\ell\right) \leq 3^d \sum_{j=1}^{k} m(B_{i_j}) &\leq \frac{3^d}{\alpha} \sum_{j=1}^{k} \int_{B_{i_j}} |f(y)|\, dy \\ &= \frac{3^d}{\alpha} \int_{\bigcup_{j=1}^{k} B_{i_j}} |f(y)|\, dy \\ &\leq \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(y)|\, dy. \end{aligned}$$

Since this inequality is true for all compact subsets K of E_α, the proof of the weak type inequality for the maximal operator is complete.

1.2 The Lebesgue differentiation theorem

The estimate obtained for the maximal function now leads to a solution of the averaging problem.

Theorem 1.3 *If f is integrable on $\mathbb{R}^d$, then*

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\, dy = f(x) \qquad \text{for a.e. } x. \tag{4}$$

Proof. It suffices to show that for each $\alpha > 0$ the set

$$E_\alpha = \left\{ x : \limsup_{\substack{m(B) \to 0 \\ x \in B}} \left| \frac{1}{m(B)} \int_B f(y)\, dy - f(x) \right| > 2\alpha \right\}$$

has measure zero, because this assertion then guarantees that the set $E = \bigcup_{n=1}^{\infty} E_{1/n}$ has measure zero, and the limit in (4) holds at all points of E^c.

We fix α, and recall Theorem 2.4 in Chapter 2, which states that for each $\epsilon > 0$ we may select a continuous function g of compact support with $\|f - g\|_{L^1(\mathbb{R}^d)} < \epsilon$. As we remarked earlier, the continuity of g implies that

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B g(y)\, dy = g(x), \qquad \text{for all } x.$$

Since we may write the difference $\frac{1}{m(B)} \int_B f(y)\, dy - f(x)$ as

$$\frac{1}{m(B)} \int_B (f(y) - g(y))\, dy + \frac{1}{m(B)} \int_B g(y)\, dy - g(x) + g(x) - f(x)$$

we find that

$$\limsup_{\substack{m(B) \to 0 \\ x \in B}} \left| \frac{1}{m(B)} \int_B f(y)\, dy - f(x) \right| \leq (f-g)^*(x) + |g(x) - f(x)|,$$

where the symbol $*$ indicates the maximal function. Consequently, if

$$F_\alpha = \{x : \ (f-g)^*(x) > \alpha\} \quad \text{and} \quad G_\alpha = \{x : \ |f(x) - g(x)| > \alpha\}$$

then $E_\alpha \subset (F_\alpha \cup G_\alpha)$, because if u_1 and u_2 are positive, then $u_1 + u_2 > 2\alpha$ only if $u_i > \alpha$ for at least one u_i. On the one hand, Tchebychev's inequality yields

$$m(G_\alpha) \leq \frac{1}{\alpha} \, \|f - g\|_{L^1(\mathbb{R}^d)},$$

and on the other hand, the weak type estimate for the maximal function gives

$$m(F_\alpha) \leq \frac{A}{\alpha} \, \|f - g\|_{L^1(\mathbb{R}^d)}.$$

The function g was selected so that $\|f - g\|_{L^1(\mathbb{R}^d)} < \epsilon$. Hence we get

$$m(E_\alpha) \leq \frac{A}{\alpha}\, \epsilon + \frac{1}{\alpha}\, \epsilon.$$

Since ϵ is arbitrary, we must have $m(E_\alpha) = 0$, and the proof of the theorem is complete.

Note that as an immediate consequence of the theorem applied to $|f|$, we see that $f^*(x) \geq |f(x)|$ for a.e. x, with f^* the maximal function.

We have worked so far under the assumption that f is integrable. This "global" assumption is slightly out of place in the context of a "local" notion like differentiability. Indeed, the limit in Lebesgue's theorem is taken over balls that shrink to the point x, so the behavior of f far from x is irrelevant. Thus, we expect the result to remain valid if we simply assume integrability of f on every ball.

To make this precise, we say that a measurable function f on $\mathbb{R}^d$ is **locally integrable**, if for every ball B the function $f(x)\chi_B(x)$ is integrable. We shall denote by $L^1_{\text{loc}}(\mathbb{R}^d)$ the space of all locally integrable functions. Loosely speaking, the behavior at infinity does not affect the local integrability of a function. For example, the functions $e^{|x|}$ and $|x|^{-1/2}$ are both locally integrable, but not integrable on $\mathbb{R}^d$.

Clearly, the conclusion of the last theorem holds under the weaker assumption that f is locally integrable.

Theorem 1.4 *If* $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, *then*

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\, dy = f(x), \qquad \textit{for a.e. } x.$$

Our first application of this theorem yields an interesting insight into the nature of measurable sets. If E is a measurable set and $x \in \mathbb{R}^d$, we say that x is a point of **Lebesgue density** of E if

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{m(B \cap E)}{m(B)} = 1.$$

Loosely speaking, this condition says that small balls around x are almost entirely covered by E. More precisely, for every $\alpha < 1$ close to 1, and every ball of sufficiently small radius containing x, we have

$$m(B \cap E) \geq \alpha m(B).$$

Thus E covers at least a proportion α of B.

An application of Theorem 1.4 to the characteristic function of E immediately yields the following:

Corollary 1.5 *Suppose* E *is a measurable subset of* $\mathbb{R}^d$. *Then:*

(i) *Almost every* $x \in E$ *is a point of density of* E.

(ii) *Almost every* $x \notin E$ *is not a point of density of* E.

We next consider a notion that for integrable functions serves as a useful substitute for pointwise continuity.

If f is locally integrable on $\mathbb{R}^d$, the **Lebesgue set** of f consists of all points $\overline{x} \in \mathbb{R}^d$ for which $f(\overline{x})$ is finite and

$$\lim_{\substack{m(B) \to 0 \\ \overline{x} \in B}} \frac{1}{m(B)} \int_B |f(y) - f(\overline{x})|\, dy = 0.$$

At this stage, two simple observations about this definition are in order. First, $\overline{x}$ belongs to the Lebesgue set of f whenever f is continuous at $\overline{x}$. Second, if $\overline{x}$ is in the Lebesgue set of f, then

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\, dy = f(\overline{x}).$$

Corollary 1.6 *If* f *is locally integrable on* $\mathbb{R}^d$, *then almost every point belongs to the Lebesgue set of* f.

Proof. An application of Theorem 1.4 to the function $|f(y) - r|$ shows that for each rational r, there exists a set E_r of measure zero, such that

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B |f(y) - r| \, dy = |f(x) - r| \quad \text{whenever } x \notin E_r.$$

If $E = \bigcup_{r \in \mathbb{Q}} E_r$, then $m(E) = 0$. Now suppose that $\overline{x} \notin E$ and $f(\overline{x})$ is finite. Given $\epsilon > 0$, there exists a rational r such that $|f(\overline{x}) - r| < \epsilon$. Since

$$\frac{1}{m(B)} \int_B |f(y) - f(\overline{x})| \, dy \leq \frac{1}{m(B)} \int_B |f(y) - r| \, dy + |f(\overline{x}) - r|,$$

we must have

$$\limsup_{\substack{m(B) \to 0 \\ \overline{x} \in B}} \frac{1}{m(B)} \int_B |f(y) - f(\overline{x})| \, dy \leq 2\epsilon,$$

and thus $\overline{x}$ is in the Lebesgue set of f. The corollary is therefore proved.

Remark. Recall from the definition in Section 2 of Chapter 2 that elements of $L^1(\mathbb{R}^d)$ are actually equivalence classes, with two functions being equivalent if they differ on a set of measure zero. It is interesting to observe that the set of points where the averages (4) converge to a limit is independent of the representation of f chosen, because

$$\int_B f(y) \, dy = \int_B g(y) \, dy$$

whenever f and g are equivalent. Nevertheless, the Lebesgue set of f depends on the particular representative of f that we consider.

We shall see that the Lebesgue set of a function enjoys a universal property in that at its points the function can be recovered by a wide variety of averages. We will prove this both for averages over sets that generalize balls, and in the setting of approximations to the identity. Note that the theory of differentiation developed so far uses averages over balls, but as we mentioned earlier, one could ask whether similar conclusions hold for other families of sets, such as cubes or rectangles. The answer depends in a fundamental way on the geometric properties of the family in question. For example, we now show that in the case of cubes (and more generally families of sets with bounded "eccentricity") the above results carry over. However, in the case of the family of *all*

rectangles the existence of the limit almost everywhere and the weak type inequality fail (see Problem 8).

A collection of sets $\{U_\alpha\}$ is said to **shrink regularly** to $\overline{x}$ (or has **bounded eccentricity** at $\overline{x}$) if there is a constant $c > 0$ such that for each U_α there is a ball B with

$$\overline{x} \in B, \quad U_\alpha \subset B, \quad \text{and} \quad m(U_\alpha) \geq c\, m(B).$$

Thus U_α is contained in B, but its measure is *comparable* to the measure of B. For example, the set of all open cubes containing $\overline{x}$ shrink regularly to $\overline{x}$. However, in $\mathbb{R}^d$ with $d \geq 2$ the collection of all open rectangles containing $\overline{x}$ does not shrink regularly to $\overline{x}$. This can be seen if we consider very thin rectangles.

Corollary 1.7 *Suppose f is locally integrable on $\mathbb{R}^d$. If $\{U_\alpha\}$ shrinks regularly to $\overline{x}$, then*

$$\lim_{\substack{m(U_\alpha) \to 0 \\ x \in U_\alpha}} \frac{1}{m(U_\alpha)} \int_{U_\alpha} f(y)\, dy = f(\overline{x})$$

for every point $\overline{x}$ in the Lebesgue set of f.

The proof is immediate once we observe that if $\overline{x} \in B$ with $U_\alpha \subset B$ and $m(U_\alpha) \geq cm(B)$, then

$$\frac{1}{m(U_\alpha)} \int_{U_\alpha} |f(y) - f(\overline{x})|\, dy \leq \frac{1}{cm(B)} \int_B |f(y) - f(\overline{x})|\, dy.$$

2 Good kernels and approximations to the identity

We shall now turn to averages of functions given as convolutions,[2] which can be written as

$$(f * K_\delta)(x) = \int_{\mathbb{R}^d} f(x-y) K_\delta(y)\, dy.$$

Here f is a general integrable function, which we keep fixed, while the K_δ vary over a specific family of functions, referred to as kernels. Expressions of this kind arise in many questions (for instance, in the Fourier inversion theorem of the previous chapter), and were already discussed in Book I.

In our initial consideration we called these functions “good kernels” if they are integrable and satisfy the following conditions for $\delta > 0$:

[2]Some basic properties of convolutions are described in Exercise 21 of the previous chapter.

(i) $\displaystyle\int_{\mathbb{R}^d} K_\delta(x)\,dx = 1.$

(ii) $\displaystyle\int_{\mathbb{R}^d} |K_\delta(x)|\,dx \leq A.$

(iii) For every $\eta > 0$,

$$\int_{|x|\geq\eta} |K_\delta(x)|\,dx \to 0 \qquad \text{as } \delta \to 0.$$

Here A is a constant independent of δ.

The main use of these kernels was that whenever f is bounded, then $(f * K_\delta)(x) \to f(x)$ as $\delta \to 0$, at every point of continuity of f. To obtain a similar conclusion, one also valid at all points of the Lebesgue set of f, we need to strengthen somewhat our assumptions on the kernels K_δ. To reflect this situation we adopt a different terminology and refer to the resulting narrower class of kernels as **approximations to the identity**. The assumptions are again that the K_δ are integrable and satisfy conditions (i) but, instead of (ii) and (iii), we assume:

(ii′) $|K_\delta(x)| \leq A\delta^{-d}$ for all $\delta > 0$.

(iii′) $|K_\delta(x)| \leq A\delta/|x|^{d+1}$ for all $\delta > 0$ and $x \in \mathbb{R}^d$.[3]

We observe that these requirements are stronger and imply the conditions in the definition of good kernels. Indeed, we first prove (ii). For that, we use the second illustration of Corollary 1.10 in Chapter 2, which gives

$$\int_{|x|\geq\epsilon} \frac{dx}{|x|^{d+1}} \leq \frac{C}{\epsilon} \qquad \text{for some } C > 0 \text{ and all } \epsilon > 0. \tag{5}$$

Then, using the estimates (ii′) and (iii′) when $|x| < \delta$ and $|x| \geq \delta$, respectively, yields

$$\begin{aligned}
\int_{\mathbb{R}^d} |K_\delta(x)|\,dx &= \int_{|x|<\delta} |K_\delta(x)|\,dx + \int_{|x|\geq\delta} |K_\delta(x)|\,dx \\
&\leq A\int_{|x|<\delta} \frac{dx}{\delta^d} + A\delta\int_{|x|\geq\delta} \frac{1}{|x|^{d+1}}\,dx \\
&\leq A' + A'' < \infty.
\end{aligned}$$

[3]Sometimes the condition (iii′) is replaced by the requirement $|K_\delta(x)| \leq A\delta^\epsilon/|x|^{d+\epsilon}$ for some fixed $\epsilon > 0$. However, the special case $\epsilon = 1$ suffices in most circumstances.

Finally, the last condition of a good kernel is also verified, since another application of (5) gives

$$\int_{|x|\geq\eta} |K_\delta(x)|\, dx \leq A\delta \int_{|x|\geq\eta} \frac{dx}{|x|^{d+1}}$$
$$\leq \frac{A'\delta}{\eta},$$

and this last expression tends to 0 as $\delta \to 0$.

The term "approximation to the identity" originates in the fact that the mapping $f \mapsto f * K_\delta$ converges to the identity mapping $f \mapsto f$, as $\delta \to 0$, in various senses, as we shall see below. It is also connected with the following heuristics. Figure 2 pictures a typical approximation to the identity: for each $\delta > 0$, the kernel is supported on the set $|x| < \delta$ and has height $1/2\delta$. As δ tends to 0, this family of kernels converges to the

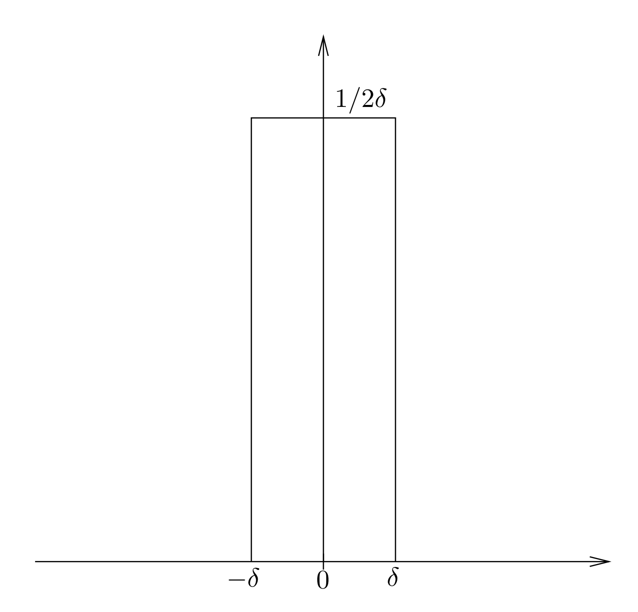


Figure 2. An approximation to the identity

so-called unit mass at the origin or **Dirac delta** "function." The latter is heuristically defined by

$$\mathcal{D}(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases} \qquad \text{and} \qquad \int \mathcal{D}(x)\, dx = 1.$$

Since each K_δ integrates to 1, we may say loosely that

$$K_\delta \to \mathcal{D} \qquad \text{as } \delta \to 0.$$

If we think of the convolution $f * \mathcal{D}$ as $\int f(x-y)\mathcal{D}(y)\,dy$, the product $f(x-y)\mathcal{D}(y)$ is 0 except when $y = 0$, and the mass of $\mathcal{D}$ is concentrated at $y = 0$, so we may intuitively expect that

$$(f * \mathcal{D})(x) = f(x).$$

Thus $f * \mathcal{D} = f$, and $\mathcal{D}$ plays the role of the identity for convolutions. We should mention that this discussion can be formalized and $\mathcal{D}$ given a precise definition either in terms of Lebesgue-Stieltjes measures, which we take up in Chapter 6, or in terms of "generalized functions" (that is, distributions), which we defer to Book IV.

We now turn to a series of examples of approximations to the identity.

EXAMPLE 1. Suppose φ is a non-negative bounded function in $\mathbb{R}^d$ that is supported on the unit ball $|x| \leq 1$, and such that

$$\int_{\mathbb{R}^d} \varphi = 1.$$

Then, if we set $K_\delta(x) = \delta^{-d}\varphi(\delta^{-1}x)$, the family $\{K_\delta\}_{\delta>0}$ is an approximation to the identity. The simple verification is left to the reader. Important special cases are in the next two examples.

EXAMPLE 2. The Poisson kernel for the upper half-plane is given by

$$\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2}, \qquad x \in \mathbb{R},$$

where the parameter is now $\delta = y > 0$.

EXAMPLE 3. The heat kernel in $\mathbb{R}^d$ is defined by

$$\mathcal{H}_t(x) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t}.$$

Here $t > 0$ and we have $\delta = t^{1/2}$. Alternatively, we could set $\delta = 4\pi t$ to make the notation consistent with the specific usage in Chapter 2.

EXAMPLE 4. The Poisson kernel for the disc is

$$\frac{1}{2\pi}P_r(x) = \begin{cases} \dfrac{1}{2\pi}\dfrac{1-r^2}{1-2r\cos x + r^2} & \text{if } |x| \leq \pi, \\ 0 & \text{if } |x| > \pi. \end{cases}$$

Here we have $0 < r < 1$ and $\delta = 1 - r$.

EXAMPLE 5. The Fejér kernel is defined by

$$\frac{1}{2\pi}F_N(x) = \begin{cases} \dfrac{1}{2\pi N}\dfrac{\sin^2(Nx/2)}{\sin^2(x/2)} & \text{if } |x| \leq \pi, \\ 0 & \text{if } |x| > \pi, \end{cases}$$

where $\delta = 1/N$.

We note that Examples 2 through 5 have already appeared in Book I.

We now turn to a general result about approximations to the identity that highlights the role of the Lebesgue set.

Theorem 2.1 *If $\{K_\delta\}_{\delta>0}$ is an approximation to the identity and f is integrable on $\mathbb{R}^d$, then*

$$(f * K_\delta)(x) \to f(x) \qquad \text{as } \delta \to 0$$

for every x in the Lebesgue set of f. In particular, the limit holds for a.e. x.

Since the integral of each kernel K_δ is equal to 1, we may write

$$(f * K_\delta)(x) - f(x) = \int [f(x-y) - f(x)]\, K_\delta(y)\, dy.$$

Consequently,

$$|(f * K_\delta)(x) - f(x)| \leq \int |f(x-y) - f(x)|\, |K_\delta(y)|\, dy,$$

and it now suffices to prove that the right-hand side tends to 0 as δ goes to 0. The argument we give depends on a simple result that we isolate in the next lemma.

Lemma 2.2 *Suppose that f is integrable on $\mathbb{R}^d$, and that x is a point of the Lebesgue set of f. Let*

$$\mathcal{A}(r) = \frac{1}{r^d}\int_{|y|\leq r} |f(x-y) - f(x)|\, dy, \qquad \text{whenever } r > 0.$$

Then $\mathcal{A}(r)$ is a continuous function of $r > 0$, and

$$\mathcal{A}(r) \to 0 \qquad \text{as } r \to 0.$$

Moreover, $\mathcal{A}(r)$ is bounded, that is, $\mathcal{A}(r) \leq M$ for some $M > 0$ and all $r > 0$.

Proof. The continuity of $\mathcal{A}(r)$ follows by invoking the absolute continuity in Proposition 1.12 of Chapter 2.

The fact that $\mathcal{A}(r)$ tends to 0 as r tends to 0 follows since x belongs to the Lebesgue set of f, and the measure of a ball of radius r is $v_d r^d$. This and the continuity of $\mathcal{A}(r)$ for $0 < r \leq 1$ show that $\mathcal{A}(r)$ is bounded when $0 < r \leq 1$. To prove that $\mathcal{A}(r)$ is bounded for $r > 1$, note that

$$\begin{aligned}\mathcal{A}(r) &\leq \frac{1}{r^d} \int_{|y|\leq r} |f(x-y)|\, dy + \frac{1}{r^d} \int_{|y|\leq r} |f(x)|\, dy \\ &\leq r^{-d} \|f\|_{L^1(\mathbb{R}^d)} + v_d |f(x)|,\end{aligned}$$

and this concludes the proof of the lemma.

We now return to the proof of the theorem. The key consists in writing the integral over $\mathbb{R}^d$ as a sum of integrals over annuli as follows:

$$\int |f(x-y) - f(x)|\, |K_\delta(y)|\, dy = \int_{|y|\leq\delta} + \sum_{k=0}^{\infty} \int_{2^k\delta < |y| \leq 2^{k+1}\delta}.$$

By using the property (ii$'$) of the approximation to the identity, the first term is estimated by

$$\begin{aligned}\int_{|y|\leq\delta} |f(x-y) - f(x)|\, |K_\delta(y)|\, dy &\leq \frac{c}{\delta^d} \int_{|y|\leq\delta} |f(x-y) - f(x)|\, dy \\ &\leq c\,\mathcal{A}(\delta).\end{aligned}$$

Each term in the sum is estimated similarly, but this time by using property (iii$'$) of approximations to the identity:

$$\begin{aligned}\int_{2^k\delta<|y|\leq 2^{k+1}\delta} &|f(x-y) - f(x)|\, |K_\delta(y)|\, dy \\ &\leq \frac{c\delta}{(2^k\delta)^{d+1}} \int_{|y|\leq 2^{k+1}\delta} |f(x-y) - f(x)|\, dy \\ &\leq \frac{c'}{2^k(2^{k+1}\delta)^d} \int_{|y|\leq 2^{k+1}\delta} |f(x-y) - f(x)|\, dy \\ &\leq c'\, 2^{-k} \mathcal{A}(2^{k+1}\delta).\end{aligned}$$

Putting these estimates together, we find that

$$|(f * K_\delta)(x) - f(x)| \leq c\,\mathcal{A}(\delta) + c' \sum_{k=0}^{\infty} 2^{-k} \mathcal{A}(2^{k+1}\delta).$$

Given $\epsilon > 0$, we first choose N so large that $\sum_{k \geq N} 2^{-k} < \epsilon$. Then, by making δ sufficiently small, we have by the lemma

$$\mathcal{A}(2^k \delta) < \epsilon/N, \qquad \text{whenever } k = 0, 1, \ldots, N-1.$$

Hence, recalling that $\mathcal{A}(r)$ is bounded, we find

$$|(f * K_\delta)(x) - f(x)| \leq C\epsilon$$

for all sufficiently small δ, and the theorem is proved.

In addition to this pointwise result, convolutions with approximations to the identity also provide convergence in the L^1-norm.

Theorem 2.3 *Suppose that f is integrable on $\mathbb{R}^d$ and that $\{K_\delta\}_{\delta>0}$ is an approximation to the identity. Then, for each $\delta > 0$, the convolution*

$$(f * K_\delta)(x) = \int_{\mathbb{R}^d} f(x-y) K_\delta(y)\, dy$$

is integrable, and

$$\|(f * K_\delta) - f\|_{L^1(\mathbb{R}^d)} \to 0, \qquad \textit{as } \delta \to 0.$$

The proof is merely a repetition in a more general context of the argument in the special case where $K_\delta(x) = \delta^{-d/2} e^{-\pi|x|^2/\delta}$ given in Section 4*, Chapter 2, and so will not be repeated.

3 Differentiability of functions

We now take up the second question raised at the beginning of this chapter, that of finding a broad condition on functions F that guarantees the identity

$$F(b) - F(a) = \int_a^b F'(x)\, dx. \tag{6}$$

There are two phenomena that make a general formulation of this identity problematic. First, because of the existence of non-differentiable functions,[4] the right-hand side of (6) might not be meaningful if we merely assumed F was continuous. Second, even if $F'(x)$ existed for every x, the function F' would not necessarily be (Lebesgue) integrable. (See Exercise 12.)

[4]In particular, there are continuous nowhere differentiable functions. See Chapter 4 in Book I, or also Chapter 7 below.

How do we deal with these difficulties? One way is by limiting ourselves to those functions F that arise as indefinite integrals (of integrable functions). This raises the issue of how to characterize such functions, and we approach that problem via the study of a wider class, the functions of bounded variation. These functions are closely related to the question of rectifiability of curves, and we start by considering this connection.

3.1 Functions of bounded variation

Let γ be a parametrized curve in the plane given by $z(t) = (x(t), y(t))$, where $a \leq t \leq b$. Here $x(t)$ and $y(t)$ are continuous real-valued functions on $[a, b]$. The curve γ is **rectifiable** if there exists $M < \infty$ such that, for any partition $a = t_0 < t_1 < \cdots < t_N = b$ of $[a, b]$,

$$\sum_{j=1}^{N} |z(t_j) - z(t_{j-1})| \leq M. \tag{7}$$

By definition, the **length** $L(\gamma)$ of the curve is the supremum over all partitions of the sum on the left-hand side, that is,

$$L(\gamma) = \sup_{a=t_0<t_1<\cdots<t_N=b} \sum_{j=1}^{N} |z(t_j) - z(t_{j-1})|.$$

Alternatively, $L(\gamma)$ is the infimum of all M that satisfy (7). Geometrically, the quantity $L(\gamma)$ is obtained by approximating the curve by polygonal lines and taking the limit of the length of these polygonal lines as the interval $[a, b]$ is partitioned more finely (see the illustration in Figure 3).

Naturally, we may now ask the following questions: What analytic condition on $x(t)$ and $y(t)$ guarantees rectifiability of the curve γ? In particular, must the derivatives of $x(t)$ and $y(t)$ exist? If so, does one have the desired formula

$$L(\gamma) = \int_a^b (x'(t)^2 + y'(t)^2)^{1/2}\, dt?$$

The answer to the first question leads directly to the class of functions of bounded variation, a class that plays a key role in the theory of differentiation.

Suppose $F(t)$ is a complex-valued function defined on $[a, b]$, and $a = t_0 < t_1 < \cdots < t_N = b$ is a partition of this interval. The variation of F

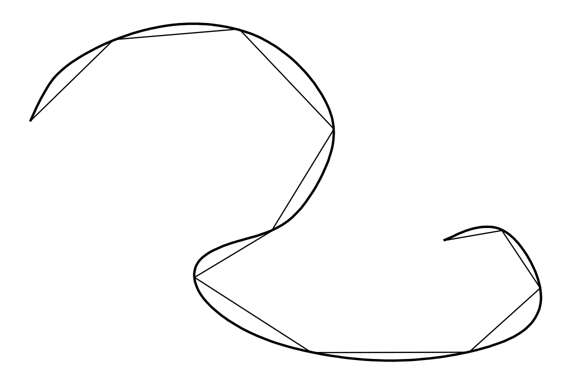


Figure 3. Approximation of a rectifiable curve by polygonal lines

on this partition is defined by

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|.$$

The function F is said to be of **bounded variation** if the variations of F over all partitions are bounded, that is, there exists $M < \infty$ so that

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| \leq M$$

for all partitions $a = t_0 < t_1 < \cdots < t_N = b$. In this definition we do not assume that F is continuous; however, when applying it to the case of curves, we will suppose that $F(t) = z(t) = x(t) + iy(t)$ is continuous.

We observe that if a partition $\tilde{\mathcal{P}}$ given by $a = \tilde{t}_0 < \tilde{t}_1 < \cdots < \tilde{t}_M = b$ is a refinement[5] of a partition $\mathcal{P}$ given by $a = t_0 < t_1 < \cdots < t_N = b$, then the variation of F on $\tilde{\mathcal{P}}$ is greater than or equal to the variation of F on $\mathcal{P}$.

Theorem 3.1 *A curve parametrized by $(x(t), y(t))$, $a \leq t \leq b$, is rectifiable if and only if both $x(t)$ and $y(t)$ are of bounded variation.*

The proof is immediate once we observe that if $F(t) = x(t) + iy(t)$, then

$$F(t_j) - F(t_{j-1}) = (x(t_j) - x(t_{j-1})) + i\,(y(t_j) - y(t_{j-1}))\,,$$

[5]We say that a partition $\tilde{\mathcal{P}}$ of $[a, b]$ is a refinement of a partition $\mathcal{P}$ of $[a, b]$ if every point in $\mathcal{P}$ also belongs to $\tilde{\mathcal{P}}$.

and if a and b are real, then $|a + ib| \leq |a| + |b| \leq 2|a + ib|$.

Intuitively, a function of bounded variation cannot oscillate too often with amplitudes that are too large. Some examples should help clarify this assertion.

We first fix some terminology. A real-valued function F defined on $[a, b]$ is **increasing** if $F(t_1) \leq F(t_2)$ whenever $a \leq t_1 \leq t_2 \leq b$. If the inequality is strict, we say that F is **strictly increasing**.

EXAMPLE 1. If F is real-valued, monotonic, and bounded, then F is of bounded variation. Indeed, if for example F is increasing and bounded by M, we see that

$$\begin{aligned}\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| &= \sum_{j=1}^{N} F(t_j) - F(t_{j-1}) \\ &= F(b) - F(a) \leq 2M.\end{aligned}$$

EXAMPLE 2. If F is differentiable at every point, and F' is bounded, then F is of bounded variation. Indeed, if $|F'| \leq M$, the mean value theorem implies

$$|F(x) - F(y)| \leq M|x - y|, \qquad \text{for all } x, y \in [a, b],$$

hence $\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| \leq M(b - a)$. (See also Exercise 23.)

EXAMPLE 3. Let

$$F(x) = \begin{cases} x^a \sin(x^{-b}) & \text{for } 0 < x \leq 1, \\ 0 & \text{if } x = 0. \end{cases}$$

Then F is of bounded variation on $[0, 1]$ if and only if $a > b$ (Exercise 11). Figure 4 illustrates the three cases $a > b$, $a = b$, and $a < b$.

The next result shows that in some sense the first example above exhausts all functions of bounded variation. For its proof, we need the following definitions. The **total variation** of f on $[a, x]$ (where $a \leq x \leq b$) is defined by

$$T_F(a, x) = \sup \sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|,$$

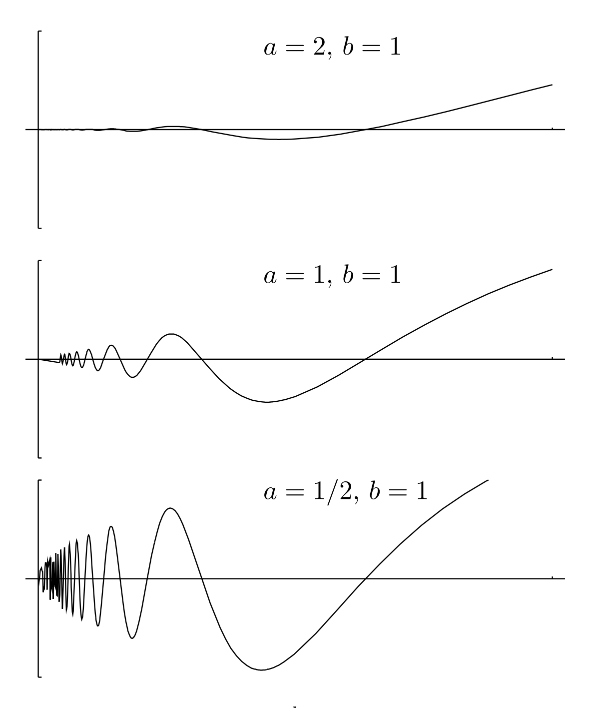


Figure 4. Graphs of $x^a \sin(x^{-b})$ for different values of a and b

where the sup is over all partitions of $[a, x]$. The preceding definition makes sense if F is complex-valued. The succeeding ones require that F is real-valued. In the spirit of the first definition, we say that the **positive variation** of F on $[a, x]$ is

$$P_F(a, x) = \sup \sum_{(+)} F(t_j) - F(t_{j-1}),$$

where the sum is over all j such that $F(t_j) \geq F(t_{j-1})$, and the supremum is over all partitions of $[a, x]$. Finally, the **negative variation** of F on $[a, x]$ is defined by

$$N_F(a, x) = \sup \sum_{(-)} -[F(t_j) - F(t_{j-1})],$$

where the sum is over all j such that $F(t_j) \leq F(t_{j-1})$, and the supremum is over all partitions of $[a, x]$.

Lemma 3.2 *Suppose F is real-valued and of bounded variation on $[a, b]$.*

Then for all $a \leq x \leq b$ one has

$$F(x) - F(a) = P_F(a,x) - N_F(a,x),$$

and

$$T_F(a,x) = P_F(a,x) + N_F(a,x).$$

Proof. Given $\epsilon > 0$ there exists a partition $a = t_0 < \cdots < t_N = x$ of $[a,x]$, such that

$$\left| P_F - \sum_{(+)} F(t_j) - F(t_{j-1}) \right| < \epsilon \text{ and } \left| N_F - \sum_{(-)} -[F(t_j) - F(t_{j-1})] \right| < \epsilon.$$

(To see this, it suffices to use the definition to obtain similar estimates for P_F and N_F with possibly different partitions, and then to consider a common refinement of these two partitions.) Since we also note that

$$F(x) - F(a) = \sum_{(+)} F(t_j) - F(t_{j-1}) - \sum_{(-)} -[F(t_j) - F(t_{j-1})],$$

we find that $|F(x) - F(a) - [P_F - N_F]| < 2\epsilon$, which proves the first identity.

For the second identity, we also note that for any partition of $a = t_0 < \cdots < t_N = x$ of $[a,x]$ we have

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| = \sum_{(+)} F(t_j) - F(t_{j-1}) + \sum_{(-)} -[F(t_j) - F(t_{j-1})],$$

hence $T_F \leq P_F + N_F$. Also, the above implies

$$\sum_{(+)} F(t_j) - F(t_{j-1}) + \sum_{(-)} -[F(t_j) - F(t_{j-1})] \leq T_F.$$

Once again, one can argue using common refinements of partitions in the definitions of P_F and N_F to deduce the inequality $P_F + N_F \leq T_F$, and the lemma is proved.

Theorem 3.3 *A real-valued function F on $[a,b]$ is of bounded variation if and only if F is the difference of two increasing bounded functions.*

Proof. Clearly, if $F = F_1 - F_2$, where each F_j is bounded and increasing, then F is of bounded variation.

Conversely, suppose F is of bounded variation. Then, we let $F_1(x) = P_F(a,x) + F(a)$ and $F_2(x) = N_F(a,x)$. Clearly, both F_1 and F_2 are increasing, of bounded variation, and by the lemma $F(x) = F_1(x) - F_2(x)$.

Observe that as a consequence, a complex-valued function of bounded variation is a (complex) linear combination of four increasing functions.

Returning to the curve γ parametrized by a continuous function $z(t) = x(t) + iy(t)$, we want to make some comment about its associated length function. Assuming that the curve is rectifiable, we define $L(A,B)$ as the length of the segment of γ that arises as the image of those t for which $A \leq t \leq B$, with $a \leq A \leq B \leq b$. Note that $L(A,B) = T_F(A,B)$, where $F(t) = z(t)$. We see that

$$L(A,C) + L(C,B) = L(A,B) \quad \text{if } A \leq C \leq B. \tag{8}$$

We also observe that $L(A,B)$ is a continuous function of B (and of A). Since it is an increasing function, to prove its continuity in B from the left, it suffices to see that for each B and $\epsilon > 0$, we can find $B_1 < B$ such that $L(A,B_1) \geq L(A,B) - \epsilon$. We do this by first finding a partition $A = t_0 < t_1 < \cdots < t_N = B$ such that the length of the corresponding polygonal line is $\geq L(A,B) - \epsilon/2$. By continuity of the function $z(t)$, we can find a B_1, with $t_{N-1} < B_1 < B$, such that $|z(B) - z(B_1)| < \epsilon/2$. Now for the refined partition $t_0 < t_1 < \cdots < t_{N-1} < B_1 < B$, the length of the polygonal line is still $\geq L(A,B) - \epsilon/2$. Therefore, the length for the partition $t_0 < t_1 < \cdots < t_{N-1} = B_1$ is $\geq L(A,B) - \epsilon$, and thus $L(A,B_1) \geq L(A,B) - \epsilon$.

To prove continuity from the right at B, let $\epsilon > 0$, pick any $C > B$, and choose a partition $B = t_0 < t_1 < \cdots < t_N = C$ such that $L(B,C) - \epsilon/2 < \sum_{j=0}^{N-1} |z(t_{j+1}) - z(t_j)|$. By considering a refinement of this partition if necessary, we may assume since z is continuous that $|z(t_1) - z(t_0)| < \epsilon/2$. If we denote $B_1 = z(t_1)$, then we get

$$L(B,C) - \epsilon/2 < \epsilon/2 + L(B_1,C).$$

Since $L(B,B_1) + L(B_1,C) = L(B,C)$ we have $L(B,B_1) < \epsilon$, and therefore $L(A,B_1) - L(A,B) < \epsilon$.

Note that what we have observed can be re-stated as follows: if a function of bounded variation is continuous, then so is its total variation.

The next result lies at the heart of the theory of differentiation.

Theorem 3.4 *If F is of bounded variation on $[a, b]$, then F is differentiable almost everywhere.*

In other words, the quotient

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h}$$

exists for almost every $x \in [a, b]$. By the previous result, it suffices to consider the case when F is increasing. In fact, we shall first also assume that F is continuous. This makes the argument simpler. As for the general case, we leave that till later. (See Section 3.3.) It will then be instructive to examine the nature of the possible discontinuities of a function of bounded variation, and reduce matters to the case of "jump functions."

We begin with a nice technical lemma of F. Riesz, which has the effect of a covering argument.

Lemma 3.5 *Suppose G is real-valued and continuous on $\mathbb{R}$. Let E be the set of points x such that*

$$G(x+h) > G(x) \quad \textit{for some } h = h_x > 0.$$

If E is non-empty, then it must be open, and hence can be written as a countable disjoint union of open intervals $E = \bigcup (a_k, b_k)$. If (a_k, b_k) is a finite interval in this union, then

$$G(b_k) - G(a_k) = 0.$$

Proof. Since G is continuous, it is clear that E is open whenever it is non-empty and can therefore be written as a disjoint union of countably many open intervals (Theorem 1.3 in Chapter 1). If (a_k, b_k) denotes a finite interval in this decomposition, then $a_k \notin E$; therefore we cannot have $G(b_k) > G(a_k)$. We now suppose that $G(b_k) < G(a_k)$. By continuity, there exists $a_k < c < b_k$ so that

$$G(c) = \frac{G(a_k) + G(b_k)}{2},$$

and in fact we may choose c farthest to the right in the interval (a_k, b_k). Since $c \in E$, there exists $d > c$ such that $G(d) > G(c)$. Since $b_k \notin E$, we must have $G(x) \leq G(b_k)$ for all $x \geq b_k$; therefore $d < b_k$. Since $G(d) > G(c)$, there exists (by continuity) $c' > d$ with $c' < b_k$ and $G(c') = G(c)$,

which contradicts the fact that c was chosen farthest to the right in (a_k, b_k). This shows that we must have $G(a_k) = G(b_k)$, and the lemma is proved.

Note. This result sometimes carries the name "rising sun lemma" for the following reason. If one thinks of the sun rising from the east (at the right) with the rays of light parallel to the x-axis, then the points $(x, G(x))$ on the graph of G, with $x \in E$, are precisely the points which are in the shade; these points appear in bold in Figure 5.

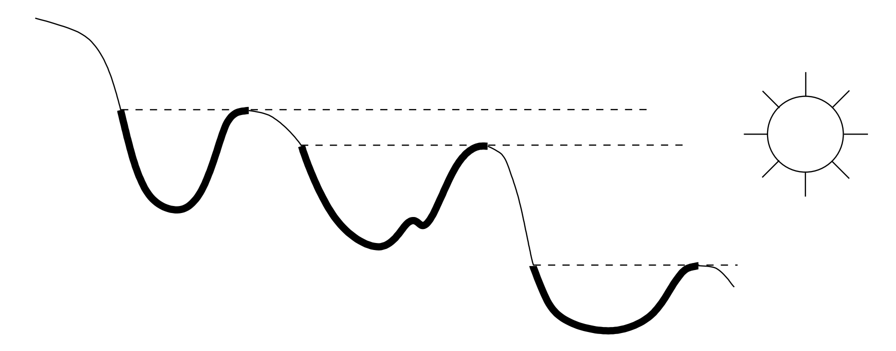


Figure 5. Rising sun lemma

A slight modification of the proof of Lemma 3.5 gives:

Corollary 3.6 *Suppose G is real-valued and continuous on a closed interval $[a, b]$. If E denotes the set of points x in (a, b) so that $G(x + h) > G(x)$ for some $h > 0$, then E is either empty or open. In the latter case, it is a disjoint union of countably many intervals (a_k, b_k), and $G(a_k) = G(b_k)$, except possibly when $a = a_k$, in which case we only have $G(a_k) \leq G(b_k)$.*

For the proof of the theorem, we define the quantity

$$\triangle_h(F)(x) = \frac{F(x + h) - F(x)}{h}.$$

We also consider the four **Dini numbers** at x defined by

$$D^+(F)(x) = \limsup_{\substack{h \to 0 \\ h > 0}} \triangle_h(F)(x)$$

$$D_+(F)(x) = \liminf_{\substack{h \to 0 \\ h > 0}} \triangle_h(F)(x)$$

$$D^-(F)(x) = \limsup_{\substack{h \to 0 \\ h < 0}} \triangle_h(F)(x)$$

$$D_-(F)(x) = \liminf_{\substack{h \to 0 \\ h < 0}} \triangle_h(F)(x).$$

Clearly, one has $D_+ \leq D^+$ and $D_- \leq D^-$. To prove the theorem it suffices to show that

(i) $D^+(F)(x) < \infty$ for a.e. x, and;

(ii) $D^+(F)(x) \leq D_-(F)(x)$ for a.e. x.

Indeed, if these results hold, then by applying (ii) to $-F(-x)$ instead of $F(x)$ we obtain $D^-(F)(x) \leq D_+(F)(x)$ for a.e. x. Therefore

$$D^+ \leq D_- \leq D^- \leq D_+ \leq D^+ < \infty \qquad \text{for a.e. } x.$$

Thus all four Dini numbers are finite and equal almost everywhere, hence $F'(x)$ exists for almost every point x.

We recall that we assume that F is increasing, bounded, and continuous on $[a, b]$. For a fixed $\gamma > 0$, let

$$E_\gamma = \{x : D^+(F)(x) > \gamma\}.$$

First, we assert that E_γ is measurable. (The proof of this simple fact is outlined in Exercise 14.) Next, we apply Corollary 3.6 to the function $G(x) = F(x) - \gamma x$, and note that we then have $E_\gamma \subset \bigcup_k (a_k, b_k)$, where $F(b_k) - F(a_k) \geq \gamma(b_k - a_k)$. Consequently,

$$\begin{aligned} m(E_\gamma) &\leq \sum_k m((a_k, b_k)) \\ &\leq \frac{1}{\gamma} \sum_k F(b_k) - F(a_k) \\ &\leq \frac{1}{\gamma}(F(b) - F(a)). \end{aligned}$$

Therefore $m(E_\gamma) \to 0$ as γ tends to infinity, and since $\{D^+F(x) < \infty\} \subset E_\gamma$ for all γ, this proves that $D^+F(x) < \infty$ almost everywhere.

Having fixed real numbers r and R such that $R > r$, we let

$$E = \{x \in [a, b] : D^+(F)(x) > R \quad \text{and} \quad r > D_-(F)(x)\}.$$

We will have shown $D^+(F)(x) \leq D_-(F)(x)$ almost everywhere once we prove that $m(E) = 0$, since it then suffices to let R and r vary over the rationals with $R > r$.

To prove that $m(E) = 0$ we may assume that $m(E) > 0$ and arrive at a contradiction. Because $R/r > 1$ we can find an open set $\mathcal{O}$ such that $E \subset \mathcal{O} \subset (a, b)$, yet $m(\mathcal{O}) < m(E) \cdot R/r$.

Now $\mathcal{O}$ can be written as $\bigcup I_n$, with I_n disjoint open intervals. Fix n and apply Corollary 3.6 to the function $G(x) = -F(-x) + rx$ on the interval $-I_n$. Reflecting through the origin again yields an open set $\bigcup_k (a_k, b_k)$ contained in I_n, where the intervals (a_k, b_k) are disjoint, with

$$F(b_k) - F(a_k) \leq r(b_k - a_k).$$

However, on each interval (a_k, b_k) we apply Corollary 3.6, this time to $G(x) = F(x) - Rx$. We thus obtain an open set $\mathcal{O}_n = \bigcup_{k,j} (a_{k,j}, b_{k,j})$ of disjoint open intervals $(a_{k,j}, b_{k,j})$ with $(a_{k,j}, b_{k,j}) \subset (a_k, b_k)$ for every j, and

$$F(b_{k,j}) - F(a_{k,j}) \geq R(b_{k,j} - a_{k,j}).$$

Then using the fact that F is increasing we find that

$$\begin{aligned} m(\mathcal{O}_n) = \sum_{k,j}(b_{k,j} - a_{k,j}) &\leq \frac{1}{R}\sum_{k,j} F(b_{k,j}) - F(a_{k,j}) \\ &\leq \frac{1}{R}\sum_{k} F(b_k) - F(a_k) \leq \frac{r}{R}\sum_{k}(b_k - a_k) \\ &\leq \frac{r}{R} m(I_n). \end{aligned}$$

Note that $\mathcal{O}_n \supset E \cap I_n$, since $D^+F(x) > R$ and $r > D_-F(x)$ for each $x \in E$; of course, $I_n \supset \mathcal{O}_n$. We now sum in n. Therefore

$$m(E) = \sum_n m(E \cap I_n) \leq \sum_n m(\mathcal{O}_n) \leq \frac{r}{R}\sum m(I_n) = \frac{r}{R} m(\mathcal{O}) < m(E).$$

The strict inequality gives a contradiction and Theorem 3.4 is proved, at least when F is continuous.

Let us see how far we have come regarding (6) if F is a monotonic function.

Corollary 3.7 *If F is increasing and continuous, then F' exists almost everywhere. Moreover F' is measurable, non-negative, and*

$$\int_a^b F'(x)\,dx \leq F(b) - F(a).$$

In particular, if F is bounded on $\mathbb{R}$, then F' is integrable on $\mathbb{R}$.

Proof. For $n \geq 1$, we consider the quotient

$$G_n(x) = \frac{F(x + 1/n) - F(x)}{1/n}.$$

By the previous theorem, we have that $G_n(x) \to F'(x)$ for a.e. x, which shows in particular that F' is measurable and non-negative.

We now extend F as a continuous function on all of $\mathbb{R}$. By Fatou's lemma (Lemma 1.7 in Chapter 2) we know that

$$\int_a^b F'(x)\,dx \leq \liminf_{n\to\infty} \int_a^b G_n(x)\,dx.$$

To complete the proof, it suffices to note that

$$\begin{aligned}
\int_a^b G_n(x)\,dx &= \frac{1}{1/n}\int_a^b F(x+1/n)\,dx - \frac{1}{1/n}\int_a^b F(x)\,dx \\
&= \frac{1}{1/n}\int_{a+1/n}^{b+1/n} F(y)\,dy - \frac{1}{1/n}\int_a^b F(x)\,dx \\
&= \frac{1}{1/n}\int_b^{b+1/n} F(x)\,dx - \frac{1}{1/n}\int_a^{a+1/n} F(x)\,dx.
\end{aligned}$$

Since F is continuous, the first and second terms converge to $F(b)$ and $F(a)$, respectively, as n goes to infinity, so the proof of the corollary is complete.

We cannot go any farther than the inequality in the corollary if we allow all continuous increasing functions, as is shown by the following important example.

The Cantor-Lebesgue function

The following simple construction yields a continuous function $F : [0,1] \to [0,1]$ that is increasing with $F(0) = 0$ and $F(1) = 1$, but $F'(x) = 0$ almost everywhere! Hence F is of bounded variation, but

$$\int_a^b F'(x)\,dx \neq F(b) - F(a).$$

Consider the standard triadic Cantor set $\mathcal{C} \subset [0,1]$ described at the end of Section 1 in Chapter 1, and recall that

$$\mathcal{C} = \bigcap_{k=0}^{\infty} C_k,$$

where each C_k is a disjoint union of 2^k closed intervals. For example, $C_1 = [0,1/3] \cup [2/3,1]$. Let $F_1(x)$ be the continuous increasing function on $[0,1]$ that satisfies $F_1(0) = 0$, $F_1(x) = 1/2$ if $1/3 \le x \le 2/3$, $F_1(1) = 1$, and F_1 is linear on C_1. Similarly, let $F_2(x)$ be continuous and increasing, and such that

$$F_2(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1/4 & \text{if } 1/9 \le x \le 2/9, \\ 1/2 & \text{if } 1/3 \le x \le 2/3, \\ 3/4 & \text{if } 7/9 \le x \le 8/9, \\ 1 & \text{if } x = 1, \end{cases}$$

and F_2 is linear on C_2. See Figure 6.

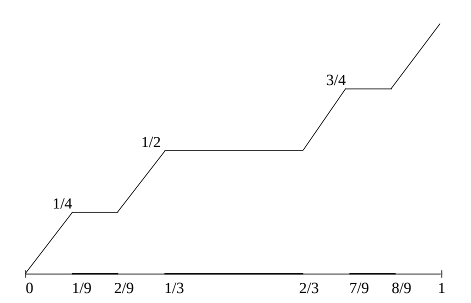


Figure 6. Construction of F_2

This process yields a sequence of continuous increasing functions $\{F_n\}_{n=1}^{\infty}$ such that clearly

$$|F_{n+1}(x) - F_n(x)| \le 2^{-n-1}.$$

Hence $\{F_n\}_{n=1}^{\infty}$ converges uniformly to a continuous limit F called the **Cantor-Lebesgue function** (Figure 7).[6] By construction, F is increasing, $F(0) = 0$, $F(1) = 1$, and we see that F is constant on each interval of the complement of the Cantor set. Since $m(\mathcal{C}) = 0$, we find that $F'(x) = 0$ almost everywhere, as desired.

[6] The reader may check that indeed this function agrees with the one given in Exercise 2 of Chapter 1.

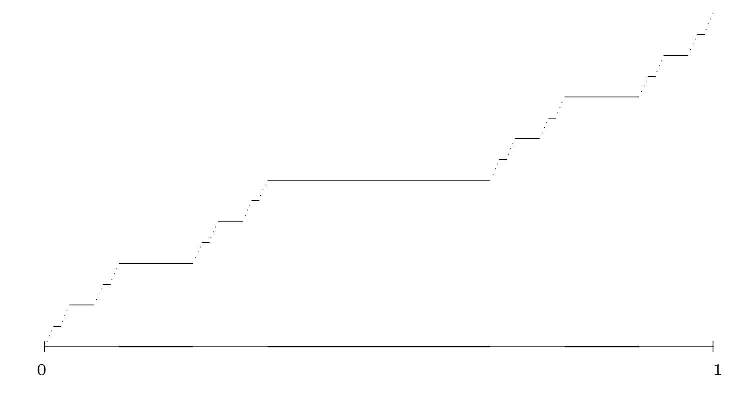


Figure 7. The Cantor-Lebesgue function

The considerations in this section, as well as this last example, show that the assumption of bounded variation guarantees the existence of a derivative almost everywhere, but not the validity of the formula

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

In the next section, we shall present a condition on a function that will completely settle the problem of establishing the above identity.

3.2 Absolutely continuous functions

A function F defined on $[a, b]$ is **absolutely continuous** if for any $\epsilon > 0$ there exists $\delta > 0$ so that

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| < \epsilon \quad \text{whenever} \quad \sum_{k=1}^{N} (b_k - a_k) < \delta,$$

and the intervals (a_k, b_k), $k = 1, \ldots, N$ are disjoint. Some general remarks are in order.

- From the definition, it is clear that absolutely continuous functions are continuous, and in fact uniformly continuous.
- If F is absolutely continuous on a bounded interval, then it is also of bounded variation on the same interval. Moreover, as is easily seen, its total variation is continuous (in fact absolutely continuous). As a consequence the decomposition of such a function F into two

monotonic functions given in Section 3.1 shows that each of these functions is continuous.

- If $F(x) = \int_a^x f(y)\,dy$ where f is integrable, then F is absolutely continuous. This follows at once from (ii) in Proposition 1.12, Chapter 2.

In fact, this last remark shows that absolute continuity is a necessary condition to impose on F if we hope to prove $\int_a^b F'(x)\,dx = F(b) - F(a)$.

Theorem 3.8 *If F is absolutely continuous on $[a, b]$, then $F'(x)$ exists almost everywhere. Moreover, if $F'(x) = 0$ for a.e. x, then F is constant.*

Since an absolutely continuous function is the difference of two continuous monotonic functions, as we have seen above, the existence of $F'(x)$ for a.e. x follows from what we have already proved. To prove that $F'(x) = 0$ a.e. implies F is constant requires a more elaborate version of the covering argument in Lemma 1.2. For the moment we revert to the generality of d dimensions to describe this.

A collection $\mathcal{B}$ of balls $\{B\}$ is said to be a **Vitali covering** of a set E if for every $x \in E$ and any $\eta > 0$ there is a ball $B \in \mathcal{B}$, such that $x \in B$ and $m(B) < \eta$. Thus every point is covered by balls of arbitrarily small measure.

Lemma 3.9 *Suppose E is a set of finite measure and $\mathcal{B}$ is a Vitali covering of E. For any $\delta > 0$ we can find finitely many balls $B_1, \ldots, B_N$ in $\mathcal{B}$ that are disjoint and so that*

$$\sum_{i=1}^{N} m(B_i) \geq m(E) - \delta.$$

Proof. We apply the elementary Lemma 1.2 iteratively, with the aim of exhausting the set E. It suffices to take δ sufficiently small, say $\delta < m(E)$, and using the just cited covering lemma, we can find an initial collection of disjoint balls $B_1, B_2, \ldots, B_{N_1}$ in $\mathcal{B}$ such that $\sum_{i=1}^{N_1} m(B_i) \geq \gamma\delta$. (For simplicity of notation, we have written $\gamma = 3^{-d}$.) Indeed, first we have $m(E') \geq \delta$ for an appropriate compact subset E' of E. Because of the compactness of E', we can cover it by finitely many balls from $\mathcal{B}$, and then the previous lemma allows us to select a disjoint sub-collection of these balls $B_1, B_2, \ldots, B_{N_1}$ such that $\sum_{i=1}^{N_1} m(B_i) \geq \gamma m(E') \geq \gamma\delta$.

With $B_1, \ldots, B_{N_1}$ as our initial sequence of balls, we consider two possibilities: either $\sum_{i=1}^{N_1} m(B_i) \geq m(E) - \delta$ and we are done with $N =$

N_1; or, contrariwise, $\sum_{i=1}^{N_1} m(B_i) < m(E) - \delta$. In the second case, with $E_2 = E - \bigcup_{i=1}^{N_1} \overline{B_i}$, we have $m(E_2) > \delta$ (recall that $m(\overline{B_i}) = m(B_i)$). We then repeat the previous argument, by choosing a compact subset E_2' of E_2 with $m(E_2') \geq \delta$, and by noting that the balls in $\mathcal{B}$ that are *disjoint* from $\bigcup_{i=1}^{N_1} \overline{B_i}$ still cover E_2 and in fact give a Vitali covering for E_2, and hence for E_2'. Thus we can choose a finite disjoint collection of these balls B_i, $N_1 < i \leq N_2$, so that $\sum_{N_1 < i \leq N_2} m(B_i) \geq \gamma\delta$. Therefore, now $\sum_{i=1}^{N_2} m(B_i) \geq 2\gamma\delta$, and the balls B_i, $1 \leq i \leq N_2$, are disjoint.

We again consider two alternatives, whether or not $\sum_{i=1}^{N_2} m(B_i) \geq m(E) - \delta$. In the first case, we are done with $N_2 = N$, and in the second case, we proceed as before. If, continuing this way, we had reached the kth stage and not stopped before then, we would have selected a collection of disjoint balls with the sum of their measures $\geq k\gamma\delta$. In any case, our process achieves the desired goal by the kth stage if $k \geq (m(E) - \delta)/\gamma\delta$, since in this case $\sum_{i=1}^{N_k} m(B_i) \geq m(E) - \delta$.

A simple consequence is the following.

Corollary 3.10 *We can arrange the choice of the balls so that*

$$m(E - \bigcup_{i=1}^{N} B_i) < 2\delta.$$

In fact, let $\mathcal{O}$ be an open set, with $\mathcal{O} \supset E$ and $m(\mathcal{O} - E) < \delta$. Since we are dealing with a Vitali covering of E, we can restrict all of our choices above to balls contained in $\mathcal{O}$. If we do this, then $(E - \bigcup_{i=1}^{N} B_i) \cup \bigcup_{i=1}^{N} B_i \subset \mathcal{O}$, where the union on the left-hand side is a disjoint union. Hence

$$m(E - \bigcup_{i=1}^{N} B_i) \leq m(\mathcal{O}) - m(\bigcup_{i=1}^{N} B_i) \leq m(E) + \delta - (m(E) - \delta) = 2\delta.$$

We now return to the situation on the real line. To complete the proof of the theorem it suffices to show that under its hypotheses we have $F(b) = F(a)$, since if that is proved, we can replace the interval $[a, b]$ by any sub-interval. Now let E be the set of those $x \in (a, b)$ where $F'(x)$ exists and is zero. By our assumption $m(E) = b - a$. Next, momentarily fix $\epsilon > 0$. Since for each $x \in E$ we have

$$\lim_{h \to 0} \left| \frac{F(x+h) - F(x)}{h} \right| = 0,$$

then for *each* $\eta > 0$ we have an open interval $I = (a_x, b_x) \subset [a, b]$ containing x, with

$$|F(b_x) - F(a_x)| \leq \epsilon(b_x - a_x) \quad \text{and } b_x - a_x < \eta.$$

The collection of these intervals forms a Vitali covering of E, and hence by the lemma, for $\delta > 0$, we can select finitely many I_i, $1 \leq i \leq N$, $I_i = (a_i, b_i)$, which are disjoint and such that

$$\sum_{i=1}^{N} m(I_i) \geq m(E) - \delta = (b - a) - \delta. \tag{9}$$

However, $|F(b_i) - F(a_i)| \leq \epsilon(b_i - a_i)$, and upon adding these inequalities we get

$$\sum_{i=1}^{N} |F(b_i) - F(a_i)| \leq \epsilon(b - a),$$

since the intervals I_i are disjoint and lie in $[a, b]$. Next consider the complement of $\bigcup_{j=1}^{N} I_j$ in $[a, b]$. It consists of finitely many closed intervals $\bigcup_{k=1}^{M}[\alpha_k, \beta_k]$ with total length $\leq \delta$ because of (9). Thus by the absolute continuity of F (if δ is chosen appropriately in terms of ϵ), $\sum_{k=1}^{M} |F(\beta_k) - F(\alpha_k)| \leq \epsilon$. Altogether, then,

$$|F(b) - F(a)| \leq \sum_{i=1}^{N} |F(b_i) - F(a_i)| + \sum_{k=1}^{M} |F(\beta_k) - F(\alpha_k)| \leq \epsilon(b - a) + \epsilon.$$

Since ϵ was positive but otherwise arbitrary, we conclude that $F(b) - F(a) = 0$, which we set out to show.

The culmination of all our efforts is contained in the next theorem. In particular, it resolves our second problem of establishing the reciprocity between differentiation and integration.

Theorem 3.11 *Suppose F is absolutely continuous on $[a, b]$. Then F' exists almost everywhere and is integrable. Moreover,*

$$F(x) - F(a) = \int_a^x F'(y)\, dy, \qquad \textit{for all } a \leq x \leq b.$$

By selecting $x = b$ we get $F(b) - F(a) = \int_a^b F'(y)\, dy$.

Conversely, if f is integrable on $[a, b]$, then there exists an absolutely continuous function F such that $F'(x) = f(x)$ almost everywhere, and in fact, we may take $F(x) = \int_a^x f(y)\, dy$.

Proof. Since we know that a real-valued absolutely continuous function is the difference of two continuous increasing functions, Corollary 3.7 shows that F' is integrable on $[a, b]$. Now let $G(x) = \int_a^x F'(y)\,dy$. Then G is absolutely continuous; hence so is the difference $G(x) - F(x)$. By the Lebesgue differentiation theorem (Theorem 1.4), we know that $G'(x) = F'(x)$ for a.e. x; hence the difference $F - G$ has derivative 0 almost everywhere. By the previous theorem we conclude that $F - G$ is constant, and evaluating this expression at $x = a$ gives the desired result.

The converse is a consequence of the observation we made earlier, namely that $\int_a^x f(y)\,dy$ is absolutely continuous, and the Lebesgue differentiation theorem, which gives $F'(x) = f(x)$ almost everywhere.

3.3 Differentiability of jump functions

We now examine monotonic functions that are not assumed to be continuous. The resulting analysis will allow us to remove the continuity assumption made earlier in the proof of Theorem 3.4.

As before, we may assume that F is increasing and bounded. In particular, these two conditions guarantee that the limits

$$F(x^-) = \lim_{\substack{y \to x \\ y < x}} F(y) \quad \text{and} \quad F(x^+) = \lim_{\substack{y \to x \\ y > x}} F(y)$$

exist. Then of course $F(x^-) \leq F(x) \leq F(x^+)$, and the function F is continuous at x if $F(x^-) = F(x^+)$; otherwise, we say that it has a jump discontinuity. Fortunately, dealing with these discontinuities is manageable, since there can only be countably many of them.

Lemma 3.12 *A bounded increasing function F on $[a, b]$ has at most countably many discontinuities.*

Proof. If F is discontinuous at x, we may choose a rational number r_x so that $F(x^-) < r_x < F(x^+)$. If f is discontinuous at x and z with $x < z$, we must have $F(x^+) \leq F(z^-)$, hence $r_x < r_z$. Consequently, to each rational number corresponds at most one discontinuity of F, hence F can have at most a countable number of discontinuities.

Now let $\{x_n\}_{n=1}^{\infty}$ denote the points where F is discontinuous, and let α_n denote the jump of F at x_n, that is, $\alpha_n = F(x_n^+) - F(x_n^-)$. Then

$$F(x_n^+) = F(x_n^-) + \alpha_n$$

and

$$F(x_n) = F(x_n^-) + \theta_n \alpha_n, \qquad \text{for some } \theta_n, \text{ with } 0 \leq \theta_n \leq 1.$$

If we let

$$j_n(x) = \begin{cases} 0 & \text{if } x < x_n, \\ \theta_n & \text{if } x = x_n, \\ 1 & \text{if } x > x_n, \end{cases}$$

then we define the **jump function** associated to F by

$$J_F(x) = \sum_{n=1}^{\infty} \alpha_n j_n(x).$$

For simplicity, and when no confusion is possible, we shall write J instead of J_F.

Our first observation is that if F is bounded, then we must have

$$\sum_{n=1}^{\infty} \alpha_n \leq F(b) - F(a) < \infty,$$

and hence the series defining J converges absolutely and uniformly.

Lemma 3.13 *If F is increasing and bounded on $[a, b]$, then:*

(i) *$J(x)$ is discontinuous precisely at the points $\{x_n\}$ and has a jump at x_n equal to that of F.*

(ii) *The difference $F(x) - J(x)$ is increasing and continuous.*

Proof. If $x \neq x_n$ for all n, each j_n is continuous at x, and since the series converges uniformly, J must be continuous at x. If $x = x_N$ for some N, then we write

$$J(x) = \sum_{n=1}^{N} \alpha_n j_n(x) + \sum_{n=N+1}^{\infty} \alpha_n j_n(x).$$

By the same argument as above, the series on the right-hand side is continuous at x. Clearly, the finite sum has a jump discontinuity at x_N of size α_N.

For (ii), we note that (i) implies at once that $F - J$ is continuous. Finally, if $y > x$ we have

$$J(y) - J(x) \leq \sum_{x < x_n \leq y} \alpha_n \leq F(y) - F(x),$$

where the last inequality follows since F is increasing. Hence

$$F(x) - J(x) \leq F(y) - J(y),$$

and the difference $F - J$ is increasing, as desired.

Since we may write $F(x) = [F(x) - J(x)] + J(x)$, our final task is to prove that J is differentiable almost everywhere.

Theorem 3.14 *If J is the jump function considered above, then $J'(x)$ exists and vanishes almost everywhere.*

Proof. Given any $\epsilon > 0$, we note that the set E of those x where

$$\limsup_{h\to 0} \frac{J(x+h) - J(x)}{h} > \epsilon \tag{10}$$

is a measurable set. (The proof of this little fact is outlined in Exercise 14 below.) Suppose $\delta = m(E)$. We need to show that $\delta = 0$. Now observe that since the series $\sum \alpha_n$ arising in the definition of J converges, then for any η, to be chosen later, we can find an N so large that $\sum_{n>N} \alpha_n < \eta$. We then write

$$J_0(x) = \sum_{n>N} \alpha_n j_n(x),$$

and because of our choice of N we have

$$J_0(b) - J_0(a) < \eta. \tag{11}$$

However, $J - J_0$ is a finite sum of terms $\alpha_n j_n(x)$, and therefore the set of points where (10) holds, with J replaced by J_0, differs from E by at most a finite set, the points $\{x_1, x_2, \ldots, x_N\}$. Thus we can find a compact set K, with $m(K) \geq \delta/2$, so that $\limsup_{h\to 0} \frac{J_0(x+h)-J_0(x)}{h} > \epsilon$ for each $x \in K$. Hence there are intervals (a_x, b_x) containing x, $x \in K$, so that $J_0(b_x) - J_0(a_x) > \epsilon(b_x - a_x)$. We can first choose a finite collection of these intervals that covers K, and then apply Lemma 1.2 to select intervals $I_1, I_2, \ldots, I_n$ which are disjoint, and for which $\sum_{j=1}^n m(I_j) \geq m(K)/3$. The intervals $I_j = (a_j, b_j)$ of course satisfy

$$J_0(b_j) - J_0(a_j) > \epsilon(b_j - a_j).$$

Now,

$$J_0(b) - J_0(a) \geq \sum_{j=1}^{N} J_0(b_j) - J_0(a_j) > \epsilon \sum (b_j - a_j) \geq \frac{\epsilon}{3} m(K) \geq \frac{\epsilon}{6}\delta.$$

Thus by (11), $\epsilon\delta/6 < \eta$, and since we are free to choose η, it follows that $\delta = 0$ and the theorem is proved.

4 Rectifiable curves and the isoperimetric inequality

We turn to the further study of rectifiable curves and take up first the validity of the formula

$$L = \int_a^b (x'(t)^2 + y'(t)^2)^{1/2}\, dt, \tag{12}$$

for the length L of the curve parametrized by $(x(t), y(t))$.

We have already seen that rectifiable curves are precisely the curves where, besides the assumed continuity of $x(t)$ and $y(t)$, these functions are of bounded variation. However a simple example shows that formula (12) does not always hold in this context. Indeed, let $x(t) = F(t)$ and $y(t) = F(t)$, where F is the Cantor-Lebesgue function and $0 \leq t \leq 1$. Then this parametrized curve traces out the straight line from $(0,0)$ to $(1,1)$ and has length $\sqrt{2}$, yet $x'(t) = y'(t) = 0$ for a.e. t.

The integral formula expressing the length of L is in fact valid if we assume in addition that the coordinate functions of the parametrization are absolutely continuous.

Theorem 4.1 *Suppose $(x(t), y(t))$ is a curve defined for $a \leq t \leq b$. If both $x(t)$ and $y(t)$ are absolutely continuous, then the curve is rectifiable, and if L denotes its length, we have*

$$L = \int_a^b (x'(t)^2 + y'(t)^2)^{1/2}\, dt.$$

Note that if $F(t) = x(t) + iy(t)$ is absolutely continuous then it is automatically of bounded variation, and hence the curve is rectifiable. The identity (12) is an immediate consequence of the proposition below, which can be viewed as a more precise version of Corollary 3.7 for absolutely continuous functions.

Proposition 4.2 *Suppose F is complex-valued and absolutely continuous on $[a, b]$. Then*

$$T_F(a, b) = \int_a^b |F'(t)|\, dt.$$

In fact, because of Theorem 3.11, for any partition $a = t_0 < t_1 < \cdots < t_N = b$ of $[a, b]$, we have

$$\begin{aligned}\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| &= \sum_{j=1}^{N} \left| \int_{t_{j-1}}^{t_j} F'(t)\, dt \right| \\ &\leq \sum_{j=1}^{N} \int_{t_{j-1}}^{t_j} |F'(t)|\, dt \\ &= \int_a^b |F'(t)|\, dt.\end{aligned}$$

So this proves

$$T_F(a, b) \leq \int_a^b |F'(t)|\, dt. \tag{13}$$

To prove the reverse inequality, fix $\epsilon > 0$, and using Theorem 2.4 in Chapter 2 find a step function g on $[a, b]$, such that $F' = g + h$ with $\int_a^b |h(t)|\, dt < \epsilon$. Set $G(x) = \int_a^x g(t)\, dt$, and $H(x) = \int_a^x h(t)\, dt$. Then $F = G + H$, and as is easily seen

$$T_F(a, b) \geq T_G(a, b) - T_H(a, b).$$

However, by (13) $T_H(a, b) < \epsilon$, so that

$$T_F(a, b) \geq T_G(a, b) - \epsilon.$$

Now partition the interval $[a, b]$, as $a = t_0 < \cdots < t_N = b$, so that the step function g is constant on each of the intervals (t_{j-1}, t_j), $j = 1, 2, \ldots, N$. Then

$$\begin{aligned}T_G(a, b) &\geq \sum_{j=1}^{N} |G(t_j) - G(t_{j-1})| \\ &= \sum_{j=1}^{N} \left| \int_{t_j - 1}^{t_j} g(t)\, dt \right| \\ &= \sum \int_{t_{j-1}}^{t_j} |g(t)|\, dt \\ &= \int_a^b |g(t)|\, dt.\end{aligned}$$

Since $\int_a^b |g(t)|\,dt \geq \int_a^b |F'(t)|\,dt - \epsilon$, we obtain as a consequence that

$$T_F(a,b) \geq \int_a^b |F'(t)|\,dt - 2\epsilon,$$

and letting $\epsilon \to 0$ we establish the assertion and also the theorem.

Now, any curve (viewed as the image of a mapping $t \mapsto z(t)$) can in fact be realized by many different parametrizations. A rectifiable curve, however, has associated to it a unique natural parametrization, the arc-length parametrization. Indeed, let $L(A,B)$ denote the length function (considered in Section 3.1), and for the variable t in $[a,b]$ set $s = s(t) = L(a,t)$. Then $s(t)$, the arc-length, is a continuous increasing function which maps $[a,b]$ to $[0,L]$, where L is the length of the curve. The **arc-length parametrization** of the curve is now given by the pair $\tilde{z}(s) = \tilde{x}(s) + i\tilde{y}(s)$, where $\tilde{z}(s) = z(t)$, for $s = s(t)$. Notice that in this way the function $\tilde{z}(s)$ is well defined on $[0,L]$, since if $s(t_1) = s(t_2)$, $t_1 < t_2$, then in fact $z(t)$ does not vary in the interval $[t_1, t_2]$ and thus $z(t_1) = z(t_2)$. Moreover $|\tilde{z}(s_1) - \tilde{z}(s_2)| \leq |s_1 - s_2|$, for all pairs $s_1, s_2 \in [0,L]$, since the left-hand side of the inequality is the distance between two points on the curve, while the right-hand side is the length of the portion of the curve joining these two points. Also, as s varies from 0 to L, $\tilde{z}(s)$ traces out the same points (in the same order) that $z(t)$ does as t varies from a to b.

Theorem 4.3 *Suppose $(x(t), y(t))$, $a \leq t \leq b$, is a rectifiable curve that has length L. Consider the arc-length parametrization $\tilde{z}(s) = (\tilde{x}(s), \tilde{y}(s))$ described above. Then $\tilde{x}$ and $\tilde{y}$ are absolutely continuous, $|\tilde{z}'(s)| = 1$ for almost every $s \in [0,L]$, and*

$$L = \int_0^L (\tilde{x}'(s)^2 + \tilde{y}'(s)^2)^{1/2}\,ds.$$

Proof. We noted that $|\tilde{z}(s_1) - \tilde{z}(s_2)| \leq |s_1 - s_2|$, so it follows immediately that $\tilde{z}(s)$ is absolutely continuous, hence differentiable almost everywhere. Moreover, this inequality also proves that $|\tilde{z}'(s)| \leq 1$, for almost every s. By definition the total variation of $\tilde{z}$ equals L, and by the previous theorem we must have $L = \int_0^L |\tilde{z}'(s)|\,ds$. Finally, we note that this identity is possible only when $|\tilde{z}'(s)| = 1$ almost everywhere.

4.1* Minkowski content of a curve

The proof we give below of the isoperimetric inequality depends in a key way on the concept of the Minkowski content. While the idea of this

content has an interest on its own right, it is particularly relevant for us here. This is because the rectifiability of a curve is tantamount to having (finite) Minkowski content, with that quantity the same as the length of the curve.

We begin our discussion of these matters with several definitions. A curve parametrized by $z(t) = (x(t), y(t))$, $a \leq t \leq b$, is said to be **simple** if the mapping $t \mapsto z(t)$ is injective for $t \in [a, b]$. It is a **closed simple curve** if the mapping $t \mapsto z(t)$ is injective for t in $[a, b)$, and $z(a) = z(b)$. More generally, a curve is **quasi-simple** if the mapping is injective for t in the complement of finitely many points in $[a, b]$.

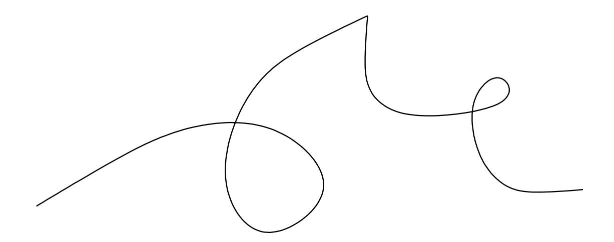


Figure 8. A quasi-simple curve

We shall find it convenient to designate by Γ the pointset traced out by the curve $z(t)$ as t varies in $[a, b]$, that is, $\Gamma = \{z(t) : a \leq t \leq b\}$. For any compact set $K \subset \mathbb{R}^2$ (we take $K = \Gamma$ below), we denote by K^δ the open set that consists of all points at distance (strictly) less than δ from K,

$$K^\delta = \{x \in \mathbb{R}^2 : d(x, K) < \delta\}.$$

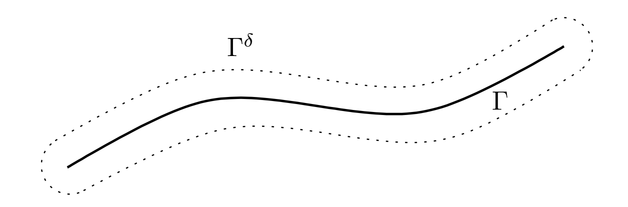


Figure 9. The curve Γ and the set Γ^δ

We then say that the set K has **Minkowski content**[7] if the limit

$$\lim_{\delta\to 0}\frac{m(K^\delta)}{2\delta}$$

exists. When this limit exists, we denote it by $\mathcal{M}(K)$.

Theorem 4.4 *Suppose* $\Gamma = \{z(t), a \leq t \leq b\}$ *is a quasi-simple curve. The Minkowski content of* Γ *exists if and only if* Γ *is rectifiable. When this is the case and* L *is the length of the curve, then* $\mathcal{M}(\Gamma) = L$.

To prove the theorem, we also consider for any compact set K

$$\mathcal{M}^*(K) = \limsup_{\delta\to 0}\frac{m(K^\delta)}{2\delta} \quad \text{and} \quad \mathcal{M}_*(K) = \liminf_{\delta\to 0}\frac{m(K^\delta)}{2\delta}$$

(both taken as extended positive numbers). Of course $\mathcal{M}_*(K) \leq \mathcal{M}^*(K)$. To say that the Minkowski content exists is the same as saying that $\mathcal{M}^*(K) < \infty$ and $\mathcal{M}_*(K) = \mathcal{M}^*(K)$. Their common value is then $\mathcal{M}(K)$.

The theorem just stated is the consequence of two propositions concerning $\mathcal{M}_*(K)$ and $\mathcal{M}^*(K)$. The first is as follows.

Proposition 4.5 *Suppose* $\Gamma = \{z(t), a \leq t \leq b\}$ *is a quasi-simple curve. If* $\mathcal{M}_*(\Gamma) < \infty$, *then the curve is rectifiable, and if* L *denotes its length, then*

$$L \leq \mathcal{M}_*(\Gamma).$$

The proof depends on the following simple observation.

Lemma 4.6 *If* $\Gamma = \{z(t), a \leq t \leq b\}$ *is any curve, and* $\Delta = |z(b) - z(a)|$ *is the distance between its end-points, then* $m(\Gamma^\delta) \geq 2\delta\Delta$.

Proof. Since the distance function and the Lebesgue measure are invariant under translations and rotations (see Section 3 in Chapter 1 and Problem 4 in Chapter 2) we may transform the situation by an appropriate composition of these motions. Therefore we may assume that the end-points of the curve have been placed on the x-axis, and thus we may suppose that $z(a) = (A, 0)$, $z(b) = (B, 0)$ with $A < B$, and $\Delta = B - A$ (in the case $A = B$ the conclusion is automatically verified).

By the continuity of the function $x(t)$, there is for each x in $[A, B]$ a value $\bar{t}$ in $[a, b]$, such that $x = x(\bar{t})$. Since $\overline{Q} = (x(\bar{t}), y(\bar{t})) \in \Gamma$, the set

[7] This is one-dimensional Minkowski content; variants are in Exercise 28 and also in Chapter 7 below.

Γ^δ contains a segment parallel to the y-axis, of length 2δ centered at $\overline{Q}$ lying above x (see Figure 10). In other words the slice $(\Gamma^\delta)_x$ contains the interval $(y(\bar{t}) - \delta, y(\bar{t}) + \delta)$, and hence $m_1((\Gamma^\delta)^x) \geq 2\delta$ (where m_1 is the one-dimensional Lebesgue measure). However by Fubini's theorem

$$m(\Gamma^\delta) = \int_{\mathbb{R}} m_1((\Gamma^\delta)_x)\, dx \geq \int_A^B m_1((\Gamma^\delta)_x)\, dx \geq 2\delta(B - A) = 2\delta\Delta,$$

and the lemma is proved.

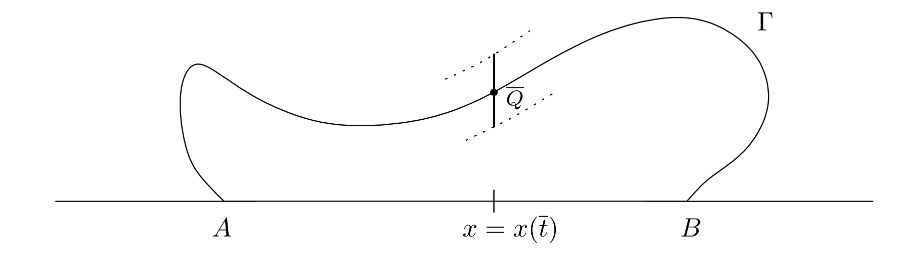


Figure 10. The situation in Lemma 4.6

We now pass to the proof of the proposition. Let us assume first that the curve is simple. Let P be any partition $a = t_0 < t_1 < \cdots < t_N = b$ of the interval $[a, b]$, and let L_P denote the length of the corresponding polygonal line, that is,

$$L_P = \sum_{j=1}^{N} |z(t_j) - z(t_{j-1})|.$$

For each $\epsilon > 0$, the continuity of $t \mapsto z(t)$ guarantees the existence of N proper closed sub-intervals $I_j = [a_j, b_j]$ of (t_{j-1}, t_j), so that

$$\sum_{j=1}^{N} |z(b_j) - z(a_j)| \geq L_P - \epsilon.$$

Let Γ_j denote the segment of the curve given by $\Gamma_j = \{z(t); t \in I_j\}$. Since the closed intervals $I_1, \ldots, I_N$ are disjoint, it follows by the simplicity of the curve that the compact sets $\Gamma_1, \Gamma_2, \ldots, \Gamma_N$ are disjoint. However, $\Gamma \supset \bigcup_{j=1}^N \Gamma_j$ and $\Gamma^\delta \supset \bigcup_{j=1}^N (\Gamma_j)^\delta$. Moreover, the disjointness of the Γ_j implies that the sets $(\Gamma_j)^\delta$ are also disjoint for sufficiently small δ. Hence

for those δ, the previous lemma applied to each Γ_j gives

$$m(\Gamma^\delta) \geq \sum_{j=1}^{N} m((\Gamma_j)^\delta) \geq 2\delta \sum |z(b_j) - z(a_j)|.$$

As a result, $m(\Gamma^\delta)/(2\delta) \geq L_P - \epsilon$, and a passage to the limit gives $\mathcal{M}_*(\Gamma) \geq L_P - \epsilon$. Since this inequality is true for all partitions P and all $\epsilon > 0$, it implies that the curve is rectifiable and its length does not exceed $\mathcal{M}_*(\Gamma)$.

The proof when the curve is merely quasi-simple is similar, except the partitions P considered must be refined so as to include as partition points those (finitely many) points in whose complement (in $[a, b]$) the mapping $t \mapsto z(t)$ is injective. The details may be left to the reader.

The second proposition is in the reverse direction.

Proposition 4.7 *Suppose* $\Gamma = \{z(t), a \leq t \leq b\}$ *is a rectifiable curve with length* L. *Then*

$$\mathcal{M}^*(\Gamma) \leq L.$$

The quantities $\mathcal{M}^*(\Gamma)$ and L are of course independent of the parametrization used; since the curve is rectifiable, it will be convenient to use the arc-length parametrization. Thus we write the curve as $z(s) = (x(s), y(s))$, with $0 \leq s \leq L$, and recall that then $z(s)$ is absolutely continuous and $|z'(s)| = 1$ for a.e. $s \in [0, L]$.

We first fix any $0 < \epsilon < 1$, and find a measurable set $E_\epsilon \subset \mathbb{R}$ and a positive number r_ϵ such that $m(E_\epsilon) < \epsilon$ and

$$\sup_{0<|h|<r_\epsilon} \left| \frac{z(s+h) - z(s)}{h} - z'(s) \right| < \epsilon \quad \text{for all } s \in [0, L] - E_\epsilon. \tag{14}$$

Indeed, for each integer n, let

$$F_n(s) = \sup_{0<|h|<1/n} \left| \frac{z(s+h) - z(s)}{h} - z'(s) \right|$$

(where $z(s)$ has been extended outside $[0, L]$, so that $z(s) = z(0)$, when $s < 0$, and $z(s) = z(L)$ when $s > L$). Because $z(s)$ is continuous the supremum of h in the definition of $F_n(s)$ can be replaced by a supremum of countably many measurable functions, and hence each F_n is measurable. However, $F_n(s) \to 0$, as $n \to \infty$ for a.e $s \in [a, b]$. Thus by Egorov's theorem the convergence is uniform outside a set E_ϵ with $m(E_\epsilon) < \epsilon$,

and so we merely need to choose $r_\epsilon = 1/n$ for sufficiently large n to establish (14). It will be convenient in what follows to assume, as we may, that $z'(s)$ exists and $|z'(s)| = 1$ for every $s \notin E_\epsilon$.

Now for any $0 < \rho < r_\epsilon$ (with $\rho < 1$), we partition the interval $[0, L]$ into consecutive closed intervals, each of length ρ, (except that the last interval may have length $\leq \rho$). Then there is a total of $N \leq L/\rho + 1$ such intervals that arise. We call these intervals $I_1, I_2, \ldots, I_N$, and divide them into two classes. The first class, those intervals I_j we call "good," are the ones that enjoy the property that $I_j \not\subset E_\epsilon$. The second class, those which are "bad," have the property that $I_j \subset E_\epsilon$. As a result, $\bigcup_{I_j \text{ bad}} I_j \subset E_\epsilon$, hence the union has measure $< \epsilon$.

We have of course that $[0, L] \subset \bigcup_{j=1}^N I_j$, and if we denote by Γ_j the segment of Γ given by $\{z(s) : s \in I_j\}$, then $\Gamma = \bigcup_{j=1}^N \Gamma_j$, and as a result $\Gamma^\delta = \bigcup_{j=1}^N (\Gamma_j)^\delta$ and $m(\Gamma^\delta) \leq \sum_{j=1}^N m((\Gamma_j)^\delta)$.

We consider first the contribution of $m((\Gamma_j)^\delta)$ when I_j is a good interval. Recall that for such $I_j = [a_j, b_j]$ there is an $s_0 \in I_j$ which is not in E_ϵ, and therefore (14) holds for $s = s_0$. Let us now visualize Γ_j by introducing a coordinate system such that $z(s_0) = 0$ and $z'(s_0) = 1$ (which we may assume after a suitable translation and rotation). We maintain the notations $z(s)$ and Γ_j for the so transformed segment of the curve.

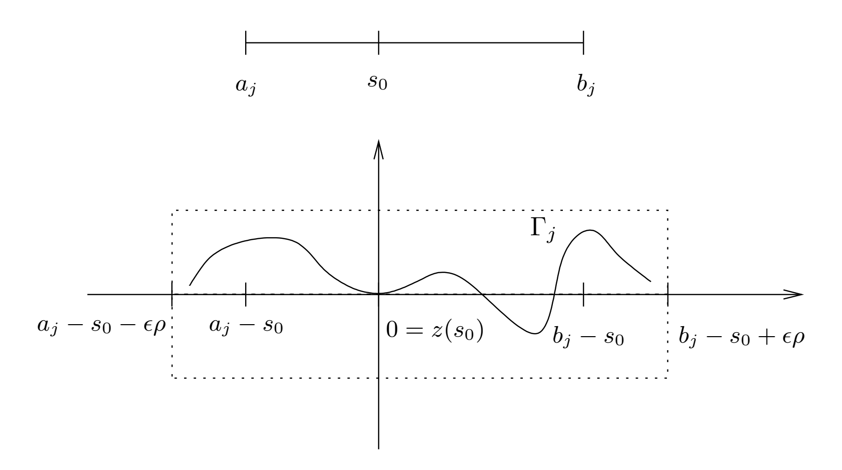


Figure 11. Estimate of $m((\Gamma_j)^\delta)$ for a good interval I_j

Note that as h varies over the interval $[a_j - s_0, b_j - s_0]$, $s_0 + h$ varies over $I_j = [a_j, b_j]$. Therefore Γ_j is contained in the rectangle

$$[a_j - s_0 - \epsilon\rho, b_j - s_0 + \epsilon\rho] \times [-\epsilon\rho, \epsilon\rho],$$

since $|h| \leq \rho < r_\epsilon$ by construction, and $|z(s_0+h) - h| < \epsilon|h|$ by (14). See Figure 11. Thus $(\Gamma_j)^\delta$ is contained in the rectangle

$$[a_j - s_0 - \epsilon\rho - \delta, b_j - s_0 + \epsilon\rho + \delta] \times [-\epsilon\rho - \delta, \epsilon\rho + \delta],$$

which has measure $\leq (\rho + 2\epsilon\rho + 2\delta)(2\epsilon\rho + 2\delta)$. Therefore, since $\epsilon \leq 1$, we have

$$m((\Gamma_j)^\delta) \leq 2\delta\rho + O(\epsilon\delta\rho + \delta^2 + \epsilon\rho^2), \tag{15}$$

where the bound arising in O is independent of ϵ, δ, and ρ. This is our desired estimate for the good intervals.

To pass to the remaining intervals we use the fact that $|z(s) - z(s')| \leq |s - s'|$ for all s and s'. Thus in every case Γ_j is contained in a ball (disc) of radius ρ, and hence $(\Gamma_j)^\delta$ is contained in a ball of radius $\rho + \delta$. Therefore we have the crude estimate

$$m((\Gamma_j)^\delta) = O(\delta^2 + \rho^2). \tag{16}$$

We now sum (15) over the good intervals (of which there are at most $L/\rho + 1$), and (16) over the bad intervals. There are at most $\epsilon/\rho + 1$ of the latter kind, since their union is included in E_ϵ and this set has measure $< \epsilon$. Altogether, then,

$$m(\Gamma^\delta) \leq 2\delta L + 2\delta\rho + O(\epsilon\delta + \delta^2/\rho + \epsilon\rho) + O\left((\epsilon/\rho + 1)(\delta^2 + \rho^2)\right),$$

which simplifies to the inequalities

$$\begin{aligned} \frac{m(\Gamma^\delta)}{2\delta} &\leq L + O\left(\rho + \epsilon + \frac{\delta}{\rho} + \frac{\epsilon\rho}{\delta} + \frac{\epsilon\delta}{\rho} + \delta + \frac{\rho^2}{\delta}\right) \\ &\leq L + O\left(\rho + \epsilon + \frac{\delta}{\rho} + \frac{\epsilon\rho}{\delta} + \frac{\rho^2}{\delta}\right), \end{aligned}$$

where in the last line we have used the fact that $\epsilon < 1$ and $\rho < 1$. In order to obtain a favorable estimate from this as $\delta \to 0$, we need to choose ρ (the length of the sub-intervals) very roughly of the same size as δ. An effective choice is $\rho = \delta/\epsilon^{1/2}$. If we fix this choice and restrict our attention to δ for which $0 < \delta < \epsilon^{1/2} r_\epsilon$, then automatically $\rho < r_\epsilon$, as required by (14). Inserting $\rho = \delta/\epsilon^{1/2}$ in the above inequality gives

$$\frac{m(\Gamma^\delta)}{2\delta} \leq L + O\left(\frac{\delta}{\epsilon^{1/2}} + \epsilon + \epsilon^{1/2} + \frac{\delta}{\epsilon}\right),$$

and thus

$$\limsup_{\delta\to 0}\frac{m(\Gamma^\delta)}{2\delta}\leq L+O(\epsilon+\epsilon^{1/2}).$$

Now we can let $\epsilon\to 0$ to obtain the desired conclusion $\mathcal{M}^*(\Gamma)\leq L$, and the proofs of the proposition and theorem are complete.

4.2* Isoperimetric inequality

The isoperimetric inequality in the plane states, in effect, that among all curves of a given length it is the circle that encloses the maximum area. A simple form of this theorem already appeared in Book I. While the proof given there had the virtue of being brief and elegant, it did suffer several shortcomings. Among them the "area" in the statement was defined indirectly via a technical artifice, and the scope of the conclusion was limited because only relatively smooth curves were considered. Here we want to remedy those defects and deal with a general version of the result.

We suppose that Ω is a bounded open subset of $\mathbb{R}^2$, and that its boundary $\overline{\Omega}-\Omega$, is a rectifiable curve Γ, with length $\ell(\Gamma)$. We do not require that Γ be a simple closed curve. The isoperimetric theorem then asserts the following.

Theorem 4.8 $4\pi\, m(\Omega)\leq\ell(\Gamma)^2$.

Proof. For each $\delta>0$ we consider the outer set

$$\Omega_+(\delta)=\{x\in\mathbb{R}^2:\ d(x,\overline{\Omega})<\delta\},$$

and the inner set

$$\Omega_-(\delta)=\{x\in\mathbb{R}^2:\ d(x,\Omega^c)\geq\delta\}.$$

Thus $\Omega_-(\delta)\subset\Omega\subset\Omega_+(\delta)$.

We notice that for $\Gamma^\delta=\{x:\ d(x,\Gamma)<\delta\}$ we have

$$\Omega_+(\delta)=\Omega_-(\delta)\cup\Gamma^\delta, \tag{17}$$

and that this union is disjoint. Moreover, if $D(\delta)$ is the open ball (disc) of radius δ centered at the origin, $D(\delta)=\{x\in\mathbb{R}^2,\ |x|<\delta\}$, then clearly

$$\begin{cases}\Omega_+(\delta) \supset \Omega+D(\delta),\\ \qquad\Omega \supset \Omega_-(\delta)+D(\delta).\end{cases} \tag{18}$$

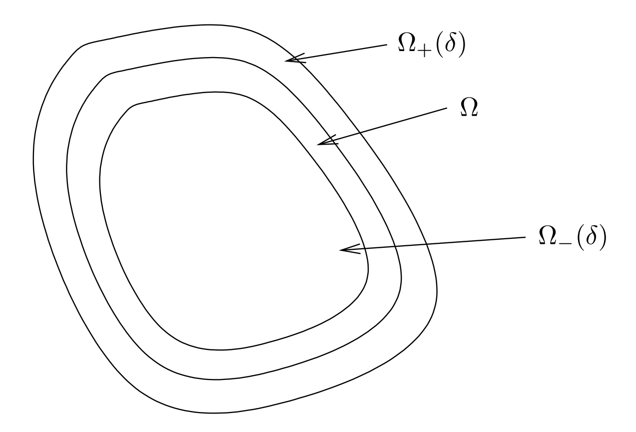


Figure 12. The sets Ω, $\Omega_-(\delta)$ and $\Omega_+(\delta)$

We now apply the Brunn-Minkowski inequality (Theorem 5.1 in Chapter 1) to the first inclusion, and obtain

$$m(\Omega_+(\delta)) \geq (m(\Omega)^{1/2} + m(D(\delta))^{1/2})^2.$$

Since $m(D(\delta)) = \pi\delta^2$ (this standard formula is established in Exercise 14 in the previous chapter), and $(A+B)^2 \geq A^2 + 2AB$ whenever A and B are positive, we find that

$$m(\Omega_+(\delta)) \geq m(\Omega) + 2\pi^{1/2}\delta\, m(\Omega)^{1/2}.$$

Similarly, $m(\Omega) \geq m(\Omega_-(\delta)) + 2\pi^{1/2}\delta\, m(\Omega_-(\delta))^{1/2}$ using the second inclusion in (18), which implies

$$-m(\Omega_-(\delta)) \geq -m(\Omega) + 2\pi^{1/2}\delta\, m(\Omega_-(\delta))^{1/2}.$$

Now by (17)

$$m(\Gamma^\delta) = m(\Omega_+(\delta)) - m(\Omega_-(\delta)),$$

and by the inequalities above, we have

$$m(\Gamma^\delta) \geq 2\pi^{1/2}\delta(m(\Omega)^{1/2} + m(\Omega_-(\delta))^{1/2}).$$

We now divide both sides by 2δ and take the limsup as $\delta \to 0$. This yields

$$\mathcal{M}^*(\Gamma) \geq \pi^{1/2}(2m(\Omega)^{1/2}),$$

since $\Omega_-(\delta) \nearrow \Omega$ as $\delta \to 0$. However, by Proposition 4.7, $\ell(\Gamma) \geq \mathcal{M}^*(\Gamma)$, so

$$\ell(\Gamma) \geq 2\pi^{1/2} m(\Omega)^{1/2},$$

which proves the theorem.

Remark. A similar result holds even without the assumption that the boundary is a (rectifiable) curve. In fact the proof shows that for any bounded open set Ω whose boundary is Γ we have

$$4\pi\, m(\Omega) \leq \mathcal{M}^*(\Gamma)^2.$$

5 Exercises

1. Suppose φ is an integrable function on $\mathbb{R}^d$ with $\int_{\mathbb{R}^d} \varphi(x)\, dx = 1$. Set $K_\delta(x) = \delta^{-d}\varphi(x/\delta)$, $\delta > 0$.

(a) Prove that $\{K_\delta\}_{\delta>0}$ is a family of good kernels.

(b) Assume in addition that φ is bounded and supported in a bounded set. Verify that $\{K_\delta\}_{\delta>0}$ is an approximation to the identity.

(c) Show that Theorem 2.3 (convergence in the L^1-norm) holds for good kernels as well.

2. Suppose $\{K_\delta\}$ is a family of kernels that satisfies:

(i) $|K_\delta(x)| \leq A\delta^{-d}$ for all $\delta > 0$.

(ii) $|K_\delta(x)| \leq A\delta/|x|^{d+1}$ for all $\delta > 0$.

(iii) $\int_{-\infty}^{\infty} K_\delta(x)\, dx = 0$ for all $\delta > 0$.

Thus K_δ satisfies conditions (i) and (ii) of approximations to the identity, but the average value of K_δ is 0 instead of 1. Show that if f is integrable on $\mathbb{R}^d$, then

$$(f * K_\delta)(x) \to 0 \qquad \text{for a.e. } x, \text{ as } \delta \to 0.$$

3. Suppose 0 is a point of (Lebesgue) density of the set $E \subset \mathbb{R}$. Show that for each of the individual conditions below there is an infinite sequence of points $x_n \in E$, with $x_n \neq 0$, and $x_n \to 0$ as $n \to \infty$.

(a) The sequence also satisfies $-x_n \in E$ for all n.

(b) In addition, $2x_n$ belongs to E for all n.

Generalize.

4. Prove that if f is integrable on $\mathbb{R}^d$, and f is not identically zero, then

$$f^*(x) \geq \frac{c}{|x|^d}, \qquad \text{for some } c > 0 \text{ and all } |x| \geq 1.$$

Conclude that f^* is not integrable on $\mathbb{R}^d$. Then, show that the weak type estimate

$$m(\{x : \ f^*(x) > \alpha\}) \leq c/\alpha$$

for all $\alpha > 0$ whenever $\int |f| = 1$, is best possible in the following sense: if f is supported in the unit ball with $\int |f| = 1$, then

$$m(\{x : \ f^*(x) > \alpha\}) \geq c'/\alpha$$

for some $c' > 0$ and all sufficiently small α.
[Hint: For the first part, use the fact that $\int_B |f| > 0$ for some ball B.]

5. Consider the function on $\mathbb{R}$ defined by

$$f(x) = \begin{cases} \dfrac{1}{|x|(\log 1/|x|)^2} & \text{if } |x| \leq 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Verify that f is integrable.

(b) Establish the inequality

$$f^*(x) \geq \frac{c}{|x|(\log 1/|x|)} \qquad \text{for some } c > 0 \text{ and all } |x| \leq 1/2,$$

to conclude that the maximal function f^* is not locally integrable.

6. In one dimension there is a version of the basic inequality (1) for the maximal function in the form of an identity. We define the "one-sided" maximal function

$$f^*_+(x) = \sup_{h>0} \frac{1}{h} \int_x^{x+h} |f(y)|\, dy.$$

If $E^+_\alpha = \{x \in \mathbb{R} : \ f^*_+(x) > \alpha\}$, then

$$m(E^+_\alpha) = \frac{1}{\alpha} \int_{E^+_\alpha} |f(y)|\, dy.$$

[Hint: Apply Lemma 3.5 to $F(x) = \int_0^x |f(y)|\, dy - \alpha x$. Then E^+_α is the union of disjoint intervals (a_k, b_k) with $\int_{a_k}^{b_k} |f(y)|\, dy = \alpha(a_k - b_k)$.]

7. Using Corollary 1.5, prove that if a measurable subset E of $[0,1]$ satisfies $m(E \cap I) \geq \alpha\, m(I)$ for some $\alpha > 0$ and all intervals I in $[0,1]$, then E has measure 1. See also Exercise 28 in Chapter 1.

8. Suppose A is a Lebesgue measurable set in $\mathbb{R}$ with $m(A) > 0$. Does there exist a sequence $\{s_n\}_{n=1}^{\infty}$ such that the complement of $\bigcup_{n=1}^{\infty}(A + s_n)$ in $\mathbb{R}$ has measure zero?

[Hint: For every $\epsilon > 0$, find an interval I_ϵ of length ℓ_ϵ such that $m(A \cap I_\epsilon) \geq (1-\epsilon)m(I_\epsilon)$. Consider $\bigcup_{k=-\infty}^{\infty}(A + t_k)$, with $t_k = k\ell_\epsilon$. Then vary ϵ.]

9. Let F be a closed subset in $\mathbb{R}$, and $\delta(x)$ the distance from x to F, that is,

$$\delta(x) = d(x, F) = \inf\{|x - y| : \; y \in F\}.$$

Clearly, $\delta(x + y) \leq |y|$ whenever $x \in F$. Prove the more refined estimate

$$\delta(x + y) = o(|y|) \qquad \text{for a.e. } x \in F,$$

that is, $\delta(x + y)/|y| \to 0$ for a.e. $x \in F$.

[Hint: Assume that x is a point of density of F.]

10. Construct an increasing function on $\mathbb{R}$ whose set of discontinuities is precisely $\mathbb{Q}$.

11. If $a, b > 0$, let

$$f(x) = \begin{cases} x^a \sin(x^{-b}) & \text{for } 0 < x \leq 1, \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is of bounded variation in $[0,1]$ if and only if $a > b$. Then, by taking $a = b$, construct (for each $0 < \alpha < 1$) a function that satisfies the Lipschitz condition of exponent α

$$|f(x) - f(y)| \leq A|x - y|^{\alpha}$$

but which is not of bounded variation.

[Hint: Note that if $h > 0$, the difference $|f(x + h) - f(x)|$ can be estimated by $C(x + h)^a$, or $C'h/x$ by the mean value theorem. Then, consider two cases, whether $x^{a+1} \geq h$ or $x^{a+1} < h$. What is the relationship between α and a?]

12. Consider the function $F(x) = x^2 \sin(1/x^2)$, $x \neq 0$, with $F(0) = 0$. Show that $F'(x)$ exists for every x, but F' is not integrable on $[-1,1]$.

13. Show directly from the definition that the Cantor-Lebesgue function is not absolutely continuous.

14. The following measurability issues arose in the discussion of differentiability of functions.

(a) Suppose F is continuous on $[a, b]$. Show that

$$D^+(F)(x) = \limsup_{\substack{h \to 0 \\ h > 0}} \frac{F(x+h) - F(x)}{h}$$

is measurable.

(b) Suppose $J(x) = \sum_{n=1}^{\infty} \alpha_n j_n(x)$ is a jump function as in Section 3.3. Show that

$$\limsup_{h \to 0} \frac{J(x+h) - J(x)}{h}$$

is measurable.

[Hint: For (a), the continuity of F allows one to restrict to countably many h in taking the limsup. For (b), given $k > m$, let $F_{k,m}^N = \sup_{1/k \leq |h| \leq 1/m} \left| \frac{J_N(x+h) - J_N(x)}{h} \right|$, where $J_N(x) = \sum_{n=1}^{N} \alpha_n j_n(x)$. Note that each $F_{k,m}^N$ is measurable. Then, successively, let $N \to \infty$, $k \to \infty$, and finally $m \to \infty$.]

15. Suppose F is of bounded variation and continuous. Prove that $F = F_1 - F_2$, where both F_1 and F_2 are monotonic and continuous.

16. Show that if F is of bounded variation in $[a, b]$, then:

(a) $\int_a^b |F'(x)|\, dx \leq T_F(a, b)$.

(b) $\int_a^b |F'(x)|\, dx = T_F(a, b)$ if and only if F is absolutely continuous.

As a result of (b), the formula $L = \int_a^b |z'(t)|\, dt$ for the length of a rectifiable curve parametrized by z holds if and only if z is absolutely continuous.

17. Prove that if $\{K_\epsilon\}_{\epsilon > 0}$ is a family of approximations to the identity, then

$$\sup_{\epsilon > 0} |(f * K_\epsilon)(x)| \leq c f^*(x)$$

for some constant $c > 0$ and all integrable f.

18. Verify the agreement between the two definitions given for the Cantor-Lebesgue function in Exercise 2, Chapter 1 and in Section 3.1 of this chapter.

19. Show that if $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous, then

(a) f maps sets of measure zero to sets of measure zero.

(b) f maps measurable sets to measurable sets.

20. This exercise deals with functions F that are absolutely continuous on $[a, b]$ and are increasing. Let $A = F(a)$ and $B = F(b)$.

(a) There exists such an F that is in addition strictly increasing, but such that $F'(x) = 0$ on a set of positive measure.

(b) The F in (a) can be chosen so that there is a measurable subset $E \subset [A, B]$, $m(E) = 0$, so that $F^{-1}(E)$ is not measurable.

(c) Prove, however, that for any increasing absolutely continuous F, and E a measurable subset of $[A, B]$, the set $F^{-1}(E) \cap \{F'(x) > 0\}$ is measurable.

[Hint: (a) Let $F(x) = \int_a^x \chi_K(x)\, dx$, where K is the complement of a Cantor-like set C of positive measure. For (b), note that $F(C)$ is a set of measure zero. Finally, for (c) prove first that $m(\mathcal{O}) = \int_{F^{-1}(\mathcal{O})} F'(x)\, dx$ for any open set $\mathcal{O}$.]

21. Let F be absolutely continuous and increasing on $[a, b]$ with $F(a) = A$ and $F(b) = B$. Suppose f is any measurable function on $[A, B]$.

(a) Show that $f(F(x))F'(x)$ is measurable on $[a, b]$. Note: $f(F(x))$ need not be measurable by Exercise 20 (b).

(b) Prove the change of variable formula: If f is integrable on $[A, B]$, then so is $f(F(x))F'(x)$, and

$$\int_A^B f(y)\, dy = \int_a^b f(F(x))F'(x)\, dx.$$

[Hint: Start with the identity $m(\mathcal{O}) = \int_{F^{-1}(\mathcal{O})} F'(x)\, dx$ used in (c) of Exercise 20 above.]

22. Suppose that F and G are absolutely continuous on $[a, b]$. Show that their product FG is also absolutely continuous. This has the following consequences.

(a) Whenever F and G are absolutely continuous in $[a, b]$,

$$\int_a^b F'(x)G(x)\, dx = -\int_a^b F(x)G'(x)\, dx + [F(x)G(x)]_a^b.$$

(b) Let F be absolutely continuous in $[-\pi, \pi]$ with $F(\pi) = F(-\pi)$. Show that if

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x)e^{-inx}\, dx,$$

such that $F(x) \sim \sum a_n e^{inx}$, then

$$F'(x) \sim \sum in a_n e^{inx}.$$

(c) What happens if $F(-\pi) \neq F(\pi)$? [Hint: Consider $F(x) = x$.]

23. Let F be continuous on $[a, b]$. Show the following.

(a) Suppose $(D^+F)(x) \geq 0$ for every $x \in [a, b]$. Then F is increasing on $[a, b]$.

(b) If $F'(x)$ exists for every $x \in (a, b)$ and $|F'(x)| \leq M$, then $|F(x) - F(y)| \leq M|x - y|$ and F is absolutely continuous.

[Hint: For (a) it suffices to show that $F(b) - F(a) \geq 0$. Assume otherwise. Hence with $G_\epsilon(x) = F(x) - F(a) + \epsilon(x - a)$, for sufficiently small $\epsilon > 0$ we have $G_\epsilon(a) = 0$, but $G_\epsilon(b) < 0$. Now let $x_0 \in [a, b)$ be the greatest value of x_0 such that $G_\epsilon(x_0) \geq 0$. However, $(D^+G_\epsilon)(x_0) > 0$.]

24. Suppose F is an increasing function on $[a, b]$.

(a) Prove that we can write

$$F = F_A + F_C + F_J,$$

where each of the functions F_A, F_C, and F_J is increasing and:

(i) F_A is absolutely continuous.

(ii) F_C is continuous, but $F'_C(x) = 0$ for a.e. x.

(iii) F_J is a jump function.

(b) Moreover, each component F_A, F_C, F_J is uniquely determined up to an additive constant.

The above is the **Lebesgue decomposition** of F. There is a corresponding decomposition for any F of bounded variation.

25. The following shows the necessity of allowing for general exceptional sets of measure zero in the differentiation Theorems 1.4, 3.4, and 3.11. Let E be any set of measure zero in $\mathbb{R}^d$. Show that:

(a) There exists a non-negative integrable f in $\mathbb{R}^d$, so that

$$\liminf_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\, dy = \infty \qquad \text{for each } x \in E.$$

(b) When $d = 1$ this may be restated as follows. There is an increasing absolutely continuous function F so that

$$D_+(F)(x) = D_-(F)(x) = \infty, \qquad \text{for each } x \in E.$$

[Hint: Find open sets $\mathcal{O}_n \supset E$, with $m(\mathcal{O}_n) < 2^{-n}$, and let $f(x) = \sum_{n=1}^{\infty} \chi_{\mathcal{O}_n}(x)$.]

26. An alternative way of defining the exterior measure $m_*(E)$ of an arbitrary set E, as given in Section 2 of Chapter 1, is to replace the coverings of E by cubes with coverings by balls. That is, suppose we define $m_*^{\mathcal{B}}(E)$ as $\inf \sum_{j=1}^{\infty} m(B_j)$, where the infimum is taken over all coverings $E \subset \bigcup_{j=1}^{\infty} B_j$ by open balls. Then $m_*(E) = m_*^{\mathcal{B}}(E)$. (Observe that this result leads to an alternate proof that the Lebesgue measure is invariant under rotations.)

Clearly $m_*(E) \leq m_*^{\mathcal{B}}(E)$. Prove the reverse inequality by showing the following. For any $\epsilon > 0$, there is a collection of balls $\{B_j\}$ such that $E \subset \bigcup_j B_j$ while $\sum_j m(B_j) \leq m_*(E) + \epsilon$. Note also that for any preassigned δ, we can choose the balls to have diameter $< \delta$.

[Hint: Assume first that E is measurable, and pick $\mathcal{O}$ open so that $\mathcal{O} \supset E$ and $m(\mathcal{O} - E) < \epsilon'$. Next, using Corollary 3.10, find balls $B_1, \ldots, B_N$ such that $\sum_{j=1}^{N} m(B_j) \leq m(E) + 2\epsilon'$ and $m(E - \bigcup_{j=1}^{N} B_j) \leq 3\epsilon'$. Finally, cover $E - \bigcup_{j=1}^{N} B_j$ by a union of cubes, the sum of whose measures is $\leq 4\epsilon'$, and replace these cubes by balls that contain them. For the general E, begin by applying the above when E is a cube.]

27. A rectifiable curve has a tangent line at almost all points of the curve. Make this statement precise.

28. A curve in $\mathbb{R}^d$ is a continuous map $t \mapsto z(t)$ of an interval $[a, b]$ into $\mathbb{R}^d$.

(a) State and prove the analogues of the conditions dealing with the rectifiability of curves and their length that are given in Theorems 3.1, 4.1, and 4.3.

(b) Define the (one-dimensional) Minkowski content $\mathcal{M}(K)$ of a compact set in $\mathbb{R}^d$ as the limit (if it exists) of

$$\frac{m(K^\delta)}{m_{d-1}(B(\delta))} \quad \text{as } \delta \to 0,$$

where $m_{d-1}(B(\delta))$ is the measure (in $\mathbb{R}^{d-1}$) of the ball defined by $B(\delta) = \{x \in \mathbb{R}^{d-1},\ |x| < \delta\}$. State and prove analogues of Propositions 4.5 and 4.7 for curves in $\mathbb{R}^d$.

29. Let $\Gamma = \{z(t),\ a \leq t \leq b\}$ be a curve, and suppose it satisfies a Lipschitz condition with exponent α, $1/2 \leq \alpha \leq 1$, that is,

$$|z(t) - z(t')| \leq A|t - t'|^\alpha \qquad \text{for all } t, t' \in [a, b].$$

Show that $m(\Gamma^\delta) = O(\delta^{2-1/\alpha})$ for $0 < \delta \leq 1$.

30. A bounded function F is said to be of bounded variation on $\mathbb{R}$ if F is of bounded variation on any finite sub-interval $[a, b]$, and $\sup_{a,b} T_F(a, b) < \infty$.

Prove that such an F enjoys the following two properties:

(a) $\int_{\mathbb{R}} |F(x+h) - F(x)|\, dx \leq A|h|$, for some constant A and all $h \in \mathbb{R}$.

(b) $|\int_{\mathbb{R}} F(x)\varphi'(x)\, dx| \leq A$, where φ ranges over all C^1 functions of bounded support with $\sup_{x\in\mathbb{R}} |\varphi(x)| \leq 1$.

For the converse, and analogues in $\mathbb{R}^d$, see Problem 6* below.

[Hint: For (a), write $F = F_1 - F_2$, where F_j are monotonic and bounded. For (b), deduce this from (a).]

31. Let F be the Cantor-Lebesgue function described in Section 3.1. Consider the curve that is the graph of F, that is, the curve given by $x(t) = t$ and $y(t) = F(t)$ with $0 \leq t \leq 1$. Prove that the length $L(\overline{x})$ of the segment $0 \leq t \leq \overline{x}$ of the curve is given by $L(\overline{x}) = \overline{x} + F(\overline{x})$. Hence the total length of the curve is 2.

32. Let $f : \mathbb{R} \to \mathbb{R}$. Prove that f satisfies the Lipschitz condition

$$|f(x) - f(y)| \leq M|x - y|$$

for some M and all $x, y \in \mathbb{R}$, if and only if f satisfies the following two properties:

(i) f is absolutely continuous.

(ii) $|f'(x)| \leq M$ for a.e. x.

6 Problems

1. Prove the following variant of the **Vitali covering lemma**: If E is covered in the Vitali sense by a family $\mathcal{B}$ of balls, and $0 < m_*(E) < \infty$, then for every $\eta > 0$ there exists a disjoint collection of balls $\{B_j\}_{j=1}^{\infty}$ in $\mathcal{B}$ such that

$$m_*\left(E/\bigcup_{j=1}^{\infty} B_j\right) = 0 \quad \text{and} \quad \sum_{j=1}^{\infty} |B_j| \leq (1+\eta)m_*(E).$$

2. The following simple one-dimensional covering lemma can be used in a number of different situations.

Suppose $I_1, I_2, \ldots, I_N$ is a given finite collection of open intervals in $\mathbb{R}$. Then there are two finite sub-collections $I_1', I_2', \ldots, I_K'$, and $I_1'', I_2'', \ldots, I_L''$, so that each sub-collection consists of mutually disjoint intervals and

$$\bigcup_{j=1}^{N} I_j = \bigcup_{k=1}^{K} I_k' \cup \bigcup_{\ell=1}^{L} I_\ell''.$$

Note that, in contrast with Lemma 1.2, the full union is covered and not merely a part.

[Hint: Choose I_1' to be an interval whose left end-point is as far left as possible. Discard all intervals contained in I_1'. If the remaining intervals are disjoint from I_1', select again an interval as far to the left as possible, and call it I_2'. Otherwise choose an interval that intersects I_1', but reaches out to the right as far as possible, and call this interval I_1''. Repeat this procedure.]

3.* There is no direct analogue of Problem 2 in higher dimensions. However, a full covering is afforded by the Besicovitch covering lemma. A version of this lemma states that there is an integer N (dependent only on the dimension d) with the following property. Suppose E is any bounded set in $\mathbb{R}^d$ that is covered by a collection $\mathcal{B}$ of balls in the (strong) sense that for each $x \in E$, there is a $B \in \mathcal{B}$ whose center is x. Then, there are N sub-collections $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_N$ of the original collection $\mathcal{B}$, such that each $\mathcal{B}_j$ is a collection of disjoint balls, and moreover,

$$E \subset \bigcup_{B \in \mathcal{B}'} B, \qquad \text{where } \mathcal{B}' = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_N.$$

4. A real-valued function φ defined on an interval (a, b) is **convex** if the region lying above its graph $\{(x, y) \in \mathbb{R}^2 : y > \varphi(x),\ a \leq x \leq b\}$ is a convex set, as defined in Section 5*, Chapter 1. Equivalently, φ is convex if

$$\varphi(\theta x_1 + (1-\theta)x_2) \leq \theta\varphi(x_1) + (1-\theta)\varphi(x_2)$$

for every $x_1, x_2 \in (a, b)$ and $0 \leq \theta \leq 1$. One can also observe as a consequence that we have the following inequality of the slopes:

$$\frac{\varphi(x+h) - \varphi(x)}{h} \leq \frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(y) - \varphi(y-h)}{h},$$

whenever $x < y$, $h > 0$, and $x + h < y$.

The following can then be proved.

(a) φ is continuous on (a, b).

(b) φ satisfies a Lipschitz condition of order 1 in any proper closed sub-interval $[a', b']$ of (a, b). Hence φ is absolutely continuous in each sub-interval.

(c) φ' exists at all but an at most denumerable number of points, and $\varphi' = D^+\varphi$ is an increasing function with

$$\varphi(y) - \varphi(x) = \int_x^y \varphi'(t)\, dt.$$

(d) Conversely, if ψ is any increasing function on (a, b), then $\varphi(x) = \int_c^x \psi(t)\, dt$ is a convex function in (a, b) (for $c \in (a, b)$).

5. Suppose that F is continuous on $[a, b]$, $F'(x)$ exists for *every* $x \in (a, b)$, and $F'(x)$ is integrable. Then F is absolutely continuous and

$$F(b) - F(a) = \int_a^b F'(x)\, dx.$$

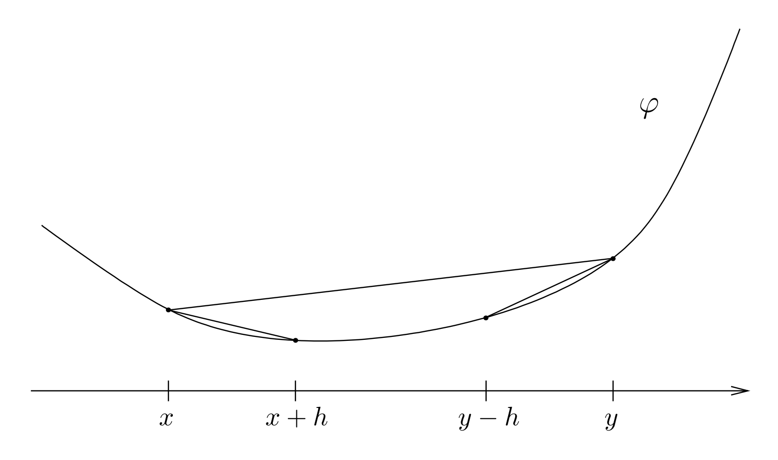


Figure 13. A convex function

[Hint: Assume $F'(x) \geq 0$ for a.e. x. We want to conclude that $F(b) \geq F(a)$. Let E be the set of measure 0 of those x such that $F'(x) < 0$. Then according to Exercise 25, there is a function Φ which is increasing, absolutely continuous, and for which $D^+\Phi(x) = \infty$, $x \in E$. Consider $F + \delta\Phi$, for each δ and apply the result (a) in Exercise 23.]

6.* The following converse to Exercise 30 characterizes functions of bounded variation.

Suppose F is a bounded measurable function on $\mathbb{R}$. If F satisfies either of conditions (a) or (b) in that exercise, then F can be modified on a set of measure zero so as to become a function of bounded variation on $\mathbb{R}$.

Moreover, on $\mathbb{R}^d$ we have the following assertion. Suppose F is a bounded measurable function on $\mathbb{R}^d$. Then the following two conditions on F are equivalent:

(a′) $\int_{\mathbb{R}^d} |F(x+h) - F(x)|\, dx \leq A|h|$, for all $h \in \mathbb{R}^d$.

(b′) $\left|\int_{\mathbb{R}^d} F(x) \frac{\partial\varphi}{\partial x_j}\, dx\right| \leq A$, for all $j = 1, \ldots, d$,

for all $\varphi \in C^1$ that have bounded support, and for which $\sup_{x\in\mathbb{R}^d} |\varphi(x)| \leq 1$.

The class of functions that satisfy either (a′) or (b′) is the extension to $\mathbb{R}^d$ of the class of functions of bounded variation.

7. Consider the function

$$f_1(x) = \sum_{n=0}^{\infty} 2^{-n} e^{2\pi i 2^n x}.$$

(a) Prove that f_1 satisfies $|f_1(x) - f_1(y)| \leq A_\alpha |x-y|^\alpha$ for each $0 < \alpha < 1$.

(b)* However, f_1 is nowhere differentiable, hence not of bounded variation.

8.* Let $\mathcal{R}$ denote the set of all rectangles in $\mathbb{R}^2$ that contain the origin, and with sides parallel to the coordinate axis. Consider the maximal operator associated to this family, namely

$$f^*_{\mathcal{R}}(x) = \sup_{R \in \mathcal{R}} \frac{1}{m(R)} \int_R |f(x-y)|\, dy.$$

(a) Then, $f \mapsto f^*_{\mathcal{R}}$ does not satisfy the weak type inequality

$$m(\{x : f^*_{\mathcal{R}}(x) > \alpha\}) \leq \frac{A}{\alpha} \|f\|_{L^1}$$

for all $\alpha > 0$, all integrable f, and some $A > 0$.

(b) Using this, one can show that there exists $f \in L^1(\mathbb{R})$ so that for $R \in \mathcal{R}$

$$\limsup_{\text{diam(R)}\to 0} \frac{1}{m(R)} \int_R f(x-y)\, dy = \infty \qquad \text{for almost every } x.$$

Here $\text{diam(R)} = \sup_{x,y \in R} |x - y|$ equals the diameter of the rectangle.

[Hint: For part (a), let B be the unit ball, and consider the function $\varphi(x) = \chi_B(x)/m(B)$. For $\delta > 0$, let $\varphi_\delta(x) = \delta^{-2}\varphi(x/\delta)$. Then

$$(\varphi_\delta)^*_{\mathcal{R}}(x) \to \frac{1}{|x_1|\,|x_2|} \qquad \text{as } \delta \to 0,$$

for every (x_1, x_2), with $x_1 x_2 \neq 0$. If the weak type inequality held, then we would have

$$m(\{|x| \leq 1 : \ |x_1 x_2|^{-1} > \alpha\}) \leq \frac{A}{\alpha}.$$

This is a contradiction since the left-hand side is of the order of $(\log \alpha)/\alpha$ as α tends to infinity.]

4 Hilbert Spaces: An Introduction

> Born barely 10 years ago, the theory of integral equations has attracted wide attention as much as for its inherent interest as for the importance of its applications. Several of its results are already classic, and no one doubts that in a few years every course in analysis will devote a chapter to it.
>
> *M. Plancherel,* 1912

There are two reasons that account for the importance of Hilbert spaces. First, they arise as the natural infinite-dimensional generalizations of Euclidean spaces, and as such, they enjoy the familiar properties of orthogonality, complemented by the important feature of completeness. Second, the theory of Hilbert spaces serves both as a conceptual framework and as a language that formulates some basic arguments in analysis in a more abstract setting.

For us the immediate link with integration theory occurs because of the example of the Lebesgue space $L^2(\mathbb{R}^d)$. The related example of $L^2([-\pi, \pi])$ is what connects Hilbert spaces with Fourier series. The latter Hilbert space can also be used in an elegant way to analyze the boundary behavior of bounded holomorphic functions in the unit disc.

A basic aspect of the theory of Hilbert spaces, as in the familiar finite-dimensional case, is the study of their linear transformations. Given the introductory nature of this chapter, we limit ourselves to rather brief discussions of several classes of such operators: unitary mappings, projections, linear functionals, and compact operators.

1 The Hilbert space L^2

A prime example of a Hilbert space is the collection of **square integrable functions** on $\mathbb{R}^d$, which is denoted by $L^2(\mathbb{R}^d)$, and consists of all complex-valued measurable functions f that satisfy

$$\int_{\mathbb{R}^d} |f(x)|^2 \, dx < \infty.$$

The resulting $L^2(\mathbb{R}^d)$**-norm** of f is defined by

$$\|f\|_{L^2(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(x)|^2\,dx\right)^{1/2}.$$

The reader should compare those definitions with these for the space $L^1(\mathbb{R}^d)$ of integrable functions and its norm that were described in Section 2, Chapter 2. A crucial difference is that L^2 has an inner product, which L^1 does not. Some relative inclusion relations between those spaces are taken up in Exercise 5.

The space $L^2(\mathbb{R}^d)$ is naturally equipped with the following inner product:

$$(f,g) = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx, \qquad \text{whenever } f,g \in L^2(\mathbb{R}^d),$$

which is intimately related to the L^2-norm since

$$(f,f)^{1/2} = \|f\|_{L^2(\mathbb{R}^d)}.$$

As in the case of integrable functions, the condition $\|f\|_{L^2(\mathbb{R}^d)} = 0$ only implies $f(x) = 0$ almost everywhere. Therefore, we in fact identify functions that are equal almost everywhere, and define $L^2(\mathbb{R}^d)$ as the space of equivalence classes under this identification. However, in practice it is often convenient to think of elements in $L^2(\mathbb{R}^d)$ as functions, and not as equivalence classes of functions.

For the definition of the inner product (f,g) to be meaningful we need to know that $f\overline{g}$ is integrable on $\mathbb{R}^d$ whenever f and g belong to $L^2(\mathbb{R}^d)$. This and other basic properties of the space of square integrable functions are gathered in the next proposition.

In the rest of this chapter we shall denote the L^2-norm by $\|\cdot\|$ (dropping the subscript $L^2(\mathbb{R}^d)$) unless stated otherwise.

Proposition 1.1 *The space $L^2(\mathbb{R}^d)$ has the following properties:*

(i) *$L^2(\mathbb{R}^d)$ is a vector space.*

(ii) *$f(x)\overline{g(x)}$ is integrable whenever $f,g \in L^2(\mathbb{R}^d)$, and the Cauchy-Schwarz inequality holds: $|(f,g)| \leq \|f\|\,\|g\|$.*

(iii) *If $g \in L^2(\mathbb{R}^d)$ is fixed, the map $f \mapsto (f,g)$ is linear in f, and also $(f,g) = \overline{(g,f)}$.*

(iv) *The triangle inequality holds: $\|f+g\| \leq \|f\| + \|g\|$.*

Proof. If $f, g \in L^2(\mathbb{R}^d)$, then since $|f(x)+g(x)| \leq 2\max(|f(x)|, |g(x)|)$, we have

$$|f(x)+g(x)|^2 \leq 4(|f(x)|^2 + |g(x)|^2),$$

therefore

$$\int |f+g|^2 \leq 4\int |f|^2 + 4\int |g|^2 < \infty,$$

hence $f+g \in L^2(\mathbb{R}^d)$. Also, if $\lambda \in \mathbb{C}$ we clearly have $\lambda f \in L^2(\mathbb{R}^d)$, and part (i) is proved.

To see why $f\overline{g}$ is integrable whenever f and g are in $L^2(\mathbb{R}^d)$, it suffices to recall that for all $A, B \geq 0$, one has $2AB \leq A^2 + B^2$, so that

$$\int |f\overline{g}| \leq \frac{1}{2}\left[\|f\|^2 + \|g\|^2\right]. \tag{1}$$

To prove the Cauchy-Schwarz inequality, we first observe that if either $\|f\| = 0$ or $\|g\| = 0$, then $fg = 0$ is zero almost everywhere, hence $(f,g) = 0$ and the inequality is obvious. Next, if we assume that $\|f\| = \|g\| = 1$, then we get the desired inequality $|(f,g)| \leq 1$. This follows from the fact that $|(f,g)| \leq \int |f\overline{g}|$, and inequality (1). Finally, in the case when both $\|f\|$ and $\|g\|$ are non-zero, we normalize f and g by setting

$$\tilde{f} = f/\|f\| \quad \text{and} \quad \tilde{g} = g/\|g\|,$$

so that $\|\tilde{f}\| = \|\tilde{g}\| = 1$. By our previous observation we then find

$$|(\tilde{f}, \tilde{g})| \leq 1.$$

Multiplying both sides of the above by $\|f\|\,\|g\|$ yields the Cauchy-Schwarz inequality.

Part (iii) follows from the linearity of the integral.

Finally, to prove the triangle inequality, we use the Cauchy-Schwarz inequality as follows:

$$\begin{aligned}\|f+g\|^2 &= (f+g, f+g) \\ &= \|f\|^2 + (f,g) + (g,f) + \|g\|^2 \\ &\leq \|f\|^2 + 2\,|(f,g)| + \|g\|^2 \\ &\leq \|f\|^2 + 2\,\|f\|\,\|g\| + \|g\|^2 \\ &= (\|f\| + \|g\|)^2,\end{aligned}$$

and taking square roots completes the argument.

We turn our attention to the notion of a limit in the space $L^2(\mathbb{R}^d)$. The norm on L^2 induces a metric d as follows: if $f, g \in L^2(\mathbb{R}^d)$, then

$$d(f, g) = \|f - g\|_{L^2(\mathbb{R}^d)}.$$

A sequence $\{f_n\} \subset L^2(\mathbb{R}^d)$ is said to be **Cauchy** if $d(f_n, f_m) \to 0$ as $n, m \to \infty$. Moreover, this sequence converges to $f \in L^2(\mathbb{R}^d)$ if $d(f_n, f) \to 0$ as $n \to \infty$.

Theorem 1.2 *The space $L^2(\mathbb{R}^d)$ is complete in its metric.*

In other words, every Cauchy sequence in $L^2(\mathbb{R}^d)$ converges to a function in $L^2(\mathbb{R}^d)$. This theorem, which is in sharp contrast with the situation for Riemann integrable functions, is a graphic illustration of the usefulness of Lebesgue's theory of integration. We elaborate on this point and its relation to Fourier series in Section 3 below.

Proof. The argument given here follows closely the proof in Chapter 2 that L^1 is complete. Let $\{f_n\}_{n=1}^\infty$ be a Cauchy sequence in L^2, and consider a subsequence $\{f_{n_k}\}_{k=1}^\infty$ of $\{f_n\}$ with the following property:

$$\|f_{n_{k+1}} - f_{n_k}\| \leq 2^{-k}, \qquad \text{for all } k \geq 1.$$

If we now consider the series whose convergence will be seen below,

$$f(x) = f_{n_1}(x) + \sum_{k=1}^\infty (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^\infty |(f_{n_{k+1}}(x) - f_{n_k}(x))|,$$

together the partial sums

$$S_K(f)(x) = f_{n_1}(x) + \sum_{k=1}^K (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$S_K(g)(x) = |f_{n_1}(x)| + \sum_{k=1}^K |f_{n_{k+1}}(x) - f_{n_k}(x)|,$$

then the triangle inequality implies

$$\begin{aligned}\|S_K(g)\| &\le \|f_{n_1}\| + \sum_{k=1}^{K} \|f_{n_{k+1}} - f_{n_k}\| \\ &\le \|f_{n_1}\| + \sum_{k=1}^{K} 2^{-k}.\end{aligned}$$

Letting K tend to infinity, and applying the monotone convergence theorem proves that $\int |g|^2 < \infty$, and since $|f| \le g$, we must have $f \in L^2(\mathbb{R}^d)$.

In particular, the series defining f converges almost everywhere, and since (by construction of the telescopic series) the $(K-1)^{\text{th}}$ partial sum of this series is precisely f_{n_K}, we find that

$$f_{n_k}(x) \to f(x) \qquad \text{a.e. } x.$$

To prove that $f_{n_k} \to f$ in $L^2(\mathbb{R}^d)$ as well, we simply observe that $|f - S_K(f)|^2 \le (2g)^2$ for all K, and apply the dominated convergence theorem to get $\|f_{n_k} - f\| \to 0$ as k tends to infinity.

Finally, the last step of the proof consists of recalling that $\{f_n\}$ is Cauchy. Given ϵ, there exists N such that for all $n, m > N$ we have $\|f_n - f_m\| < \epsilon/2$. If n_k is chosen so that $n_k > N$, and $\|f_{n_k} - f\| < \epsilon/2$, then the triangle inequality implies

$$\|f_n - f\| \le \|f_n - f_{n_k}\| + \|f_{n_k} - f\| < \epsilon$$

whenever $n > N$. This concludes the proof of the theorem.

An additional useful property of $L^2(\mathbb{R}^d)$ is contained in the following theorem.

Theorem 1.3 *The space $L^2(\mathbb{R}^d)$ is **separable**, in the sense that there exists a countable collection $\{f_k\}$ of elements in $L^2(\mathbb{R}^d)$ such that their linear combinations are dense in $L^2(\mathbb{R}^d)$.*

Proof. Consider the family of functions of the form $r\chi_R(x)$, where r is a complex number with rational real and imaginary parts, and R is a rectangle in $\mathbb{R}^d$ with rational coordinates. We claim that finite linear combinations of these type of functions are dense in $L^2(\mathbb{R}^d)$.

Suppose $f \in L^2(\mathbb{R}^d)$ and let $\epsilon > 0$. Consider for each $n \ge 1$ the function g_n defined by

$$g_n(x) = \begin{cases} f(x) & \text{if } |x| \le n \text{ and } |f(x)| \le n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $|f - g_n|^2 \leq 4|f|^2$ and $g_n(x) \to f(x)$ almost everywhere.[1] The dominated convergence theorem implies that $\|f - g_n\|^2_{L^2(\mathbb{R}^d)} \to 0$ as n tends to infinity; therefore we have

$$\|f - g_N\|_{L^2(\mathbb{R}^d)} < \epsilon/2 \quad \text{for some } N.$$

Let $g = g_N$, and note that g is a bounded function supported on a bounded set; thus $g \in L^1(\mathbb{R}^d)$. We may now find a step function φ so that $|\varphi| \leq N$ and $\int |g - \varphi| < \epsilon^2/16N$ (Theorem 2.4, Chapter 2). By replacing the coefficients and rectangles that appear in the canonical form of φ by complex numbers with rational real and imaginary parts, and rectangles with rational coordinates, we may find a ψ with $|\psi| \leq N$ and $\int |g - \psi| < \epsilon^2/8N$. Finally, we note that

$$\int |g - \psi|^2 \leq 2N \int |g - \psi| < \epsilon^2/4.$$

Consequently $\|g - \psi\| < \epsilon/2$, therefore $\|f - \psi\| < \epsilon$, and the proof is complete.

The example $L^2(\mathbb{R}^d)$ possesses all the characteristic properties of a Hilbert space, and motivates the definition of the abstract version of this concept.

2 Hilbert spaces

A set $\mathcal{H}$ is a **Hilbert space** if it satisfies the following:

(i) $\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$).[2]

(ii) $\mathcal{H}$ is equipped with an inner product $(\cdot, \cdot)$, so that

- $f \mapsto (f, g)$ is linear on $\mathcal{H}$ for every fixed $g \in \mathcal{H}$,
- $(f, g) = \overline{(g, f)}$,
- $(f, f) \geq 0$ for all $f \in \mathcal{H}$.

We let $\|f\| = (f, f)^{1/2}$.

(iii) $\|f\| = 0$ if and only if $f = 0$.

[1] By definition $f \in L^2(\mathbb{R}^d)$ implies that $|f|^2$ is integrable, hence $f(x)$ is finite for a.e x.

[2] At this stage we consider both cases, where the scalar field can be either $\mathbb{C}$ or $\mathbb{R}$. However, in many applications, such as in the context of Fourier analysis, one deals primarily with Hilbert spaces over $\mathbb{C}$.

(iv) The Cauchy-Schwarz and triangle inequalities hold

$$|(f,g)| \leq \|f\|\,\|g\| \quad \text{and} \quad \|f+g\| \leq \|f\| + \|g\|$$

for all $f, g \in \mathcal{H}$.

(v) $\mathcal{H}$ is complete in the metric $d(f,g) = \|f-g\|$.

(vi) $\mathcal{H}$ is separable.

We make two comments about the definition of a Hilbert space. First, the Cauchy-Schwarz and triangle inequalities in (iv) are in fact easy consequences of assumptions (i) and (ii). (See Exercise 1.) Second, we make the requirement that $\mathcal{H}$ be separable because that is the case in most applications encountered. That is not to say that there are no interesting non-separable examples; one such example is described in Problem 2.

Also, we remark that in the context of a Hilbert space we shall often write $\lim_{n\to\infty} f_n = f$ or $f_n \to f$ to mean that $\lim_{n\to\infty} \|f_n - f\| = 0$, which is the same as $d(f_n, f) \to 0$.

We give some examples of Hilbert spaces.

EXAMPLE 1. If E is a measurable subset of $\mathbb{R}^d$ with $m(E) > 0$, we let $L^2(E)$ denote the space of square integrable functions that are supported on E,

$$L^2(E) = \left\{ f \text{ supported on } E, \text{ so that } \int_E |f(x)|^2\,dx < \infty \right\}.$$

The inner product and norm on $L^2(E)$ are then

$$(f,g) = \int_E f(x)\overline{g(x)}\,dx \quad \text{and} \quad \|f\| = \left(\int_E |f(x)|^2\,dx \right)^{1/2}.$$

Once again, we consider two elements of $L^2(E)$ to be equivalent if they differ only on a set of measure zero; this guarantees that $\|f\| = 0$ implies $f = 0$. The properties (i) through (vi) follow from these of $L^2(\mathbb{R}^d)$ proved above.

EXAMPLE 2. A simple example is the finite-dimensional complex Euclidean space. Indeed,

$$\mathbb{C}^N = \{(a_1, \ldots, a_N) : \; a_k \in \mathbb{C}\}$$

becomes a Hilbert space when equipped with the inner product

$$\sum_{k=1}^{N} a_k \overline{b_k},$$

where $a = (a_1, \ldots, a_N)$ and $b = (b_1, \ldots, b_N)$ are in $\mathbb{C}^N$. The norm is then

$$\|a\| = \left(\sum_{k=1}^{N} |a_k|^2 \right)^{1/2}.$$

One can formulate in the same way the real Hilbert space $\mathbb{R}^N$.

Example 3. An infinite-dimensional analogue of the above example is the space $\ell^2(\mathbb{Z})$. By definition

$$\ell^2(\mathbb{Z}) = \left\{ (\ldots, a_{-2}, a_{-1}, a_0, a_1, \ldots) : \; a_i \in \mathbb{C}, \quad \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty \right\}.$$

If we denote infinite sequences by a and b, the inner product and norm on $\ell^2(\mathbb{Z})$ are

$$(a, b) = \sum_{k=-\infty}^{\infty} a_k \overline{b_k} \quad \text{and} \quad \|a\| = \left(\sum_{k=-\infty}^{\infty} |a_k|^2 \right)^{1/2}.$$

We leave the proof that $\ell^2(\mathbb{Z})$ is a Hilbert space as Exercise 4.

While this example is very simple, it will turn out that all infinite-dimensional (separable) Hilbert spaces are $\ell^2(\mathbb{Z})$ in disguise.

Also, a slight variant of this space is $\ell^2(\mathbb{N})$, where we take only one-sided sequences, that is,

$$\ell^2(\mathbb{N}) = \left\{ (a_1, a_2, \ldots) : \; a_i \in \mathbb{C}, \quad \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

The inner product and norm are then defined in the same way with the sums extending from $n = 1$ to ∞.

A characteristic feature of a Hilbert space is the notion of orthogonality. This aspect, with its rich geometric and analytic consequences, distinguishes Hilbert spaces from other normed vector spaces. We now describe some of these properties.

2.1 Orthogonality

Two elements f and g in a Hilbert space $\mathcal{H}$ with inner product $(\cdot, \cdot)$ are **orthogonal** or **perpendicular** if

$$(f, g) = 0, \qquad \text{and we then write } f \perp g.$$

The first simple observation is that the usual theorem of Pythagoras holds in the setting of abstract Hilbert spaces:

Proposition 2.1 *If $f \perp g$, then $\|f + g\|^2 = \|f\|^2 + \|g\|^2$.*

Proof. It suffices to note that $(f, g) = 0$ implies $(g, f) = 0$, and therefore

$$\begin{aligned}\|f + g\|^2 &= (f + g, f + g) = \|f\|^2 + (f, g) + (g, f) + \|g\|^2 \\ &= \|f\|^2 + \|g\|^2.\end{aligned}$$

A finite or countably infinite subset $\{e_1, e_2, \ldots\}$ of a Hilbert space $\mathcal{H}$ is **orthonormal** if

$$(e_k, e_\ell) = \begin{cases} 1 & \text{when } k = \ell, \\ 0 & \text{when } k \neq \ell. \end{cases}$$

In other words, each e_k has unit norm and is orthogonal to e_ℓ whenever $\ell \neq k$.

Proposition 2.2 *If $\{e_k\}_{k=1}^\infty$ is orthonormal, and $f = \sum a_k e_k \in \mathcal{H}$ where the sum is finite, then*

$$\|f\|^2 = \sum |a_k|^2.$$

The proof is a simple application of the Pythagorean theorem.

Given an orthonormal subset $\{e_1, e_2, \ldots\} = \{e_k\}_{k=1}^\infty$ of $\mathcal{H}$, a natural problem is to determine whether this subset spans all of $\mathcal{H}$, that is, whether finite linear combinations of elements in $\{e_1, e_2, \ldots\}$ are dense in $\mathcal{H}$. If this is the case, we say that $\{e_k\}_{k=1}^\infty$ is an **orthonormal basis** for $\mathcal{H}$. If we are in the presence of an orthonormal basis, we might expect that any $f \in \mathcal{H}$ takes the form

$$f = \sum_{k=1}^\infty a_k e_k,$$

for some constants $a_k \in \mathbb{C}$. In fact, taking the inner product of both sides with e_j, and recalling that $\{e_k\}$ is orthonormal yields (formally)

$$(f, e_j) = a_j.$$

This question is motivated by Fourier series. In fact, a good insight into the theorem below is afforded by considering the case where $\mathcal{H}$ is $L^2([-\pi, \pi])$ with inner product $(f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$, and the orthonormal set $\{e_k\}_{k=1}^{\infty}$ is merely a relabeling of the exponentials $\{e^{inx}\}_{n=-\infty}^{\infty}$.

Adapting the notation used in Fourier series, we write $f \sim \sum_{k=1}^{\infty} a_k e_k$, where $a_j = (f, e_j)$ for all j.

In the next theorem, we provide four equivalent characterizations that $\{e_k\}$ is an orthonormal basis for $\mathcal{H}$.

Theorem 2.3 *The following properties of an orthonormal set $\{e_k\}_{k=1}^{\infty}$ are equivalent.*

(i) *Finite linear combinations of elements in $\{e_k\}$ are dense in $\mathcal{H}$.*

(ii) *If $f \in \mathcal{H}$ and $(f, e_j) = 0$ for all j, then $f = 0$.*

(iii) *If $f \in \mathcal{H}$, and $S_N(f) = \sum_{k=1}^{N} a_k e_k$, where $a_k = (f, e_k)$, then $S_N(f) \to f$ as $N \to \infty$ in the norm.*

(iv) *If $a_k = (f, e_k)$, then $\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2$.*

Proof. We prove that each property implies the next, with the last one implying the first.

We begin by assuming (i). Given $f \in \mathcal{H}$ with $(f, e_j) = 0$ for all j, we wish to prove that $f = 0$. By assumption, there exists a sequence $\{g_n\}$ of elements in $\mathcal{H}$ that are finite linear combinations of elements in $\{e_k\}$, and such that $\|f - g_n\|$ tends to 0 as n goes to infinity. Since $(f, e_j) = 0$ for all j, we must have $(f, g_n) = 0$ for all n; therefore an application of the Cauchy-Schwarz inequality gives

$$\|f\|^2 = (f, f) = (f, f - g_n) \leq \|f\|\,\|f - g_n\| \quad \text{for all } n.$$

Letting $n \to \infty$ proves that $\|f\|^2 = 0$; hence $f = 0$, and (i) implies (ii).

Now suppose that (ii) is verified. For $f \in \mathcal{H}$ we define

$$S_N(f) = \sum_{k=1}^{N} a_k e_k, \qquad \text{where } a_k = (f, e_k),$$

and prove first that $S_N(f)$ converges to some element $g \in \mathcal{H}$. Indeed, one notices that the definition of a_k implies $(f - S_N(f)) \perp S_N(f)$, so the Pythagorean theorem and Proposition 2.2 give

$$\|f\|^2 = \|f - S_N(f)\|^2 + \|S_N(f)\|^2 = \|f - S_N(f)\|^2 + \sum_{k=1}^{N} |a_k|^2. \tag{2}$$

Hence $\|f\|^2 \geq \sum_{k=1}^{N} |a_k|^2$, and letting N tend to infinity we obtain **Bessel's inequality**

$$\sum_{k=1}^{\infty} |a_k|^2 \leq \|f\|^2,$$

which implies that the series $\sum_{k=1}^{\infty} |a_k|^2$ converges. Therefore, $\{S_N(f)\}_{N=1}^{\infty}$ forms a Cauchy sequence in $\mathcal{H}$ since

$$\|S_N(f) - S_M(f)\|^2 = \sum_{k=M+1}^{N} |a_k|^2 \quad \text{whenever } N > M.$$

Since $\mathcal{H}$ is complete, there exists $g \in \mathcal{H}$ such that $S_N(f) \to g$ as N tends to infinity.

Fix j, and note that for all sufficiently large N, $(f - S_N(f), e_j) = a_j - a_j = 0$. Since $S_N(f)$ tends to g, we conclude that

$$(f - g, e_j) = 0 \quad \text{for all } j.$$

Hence $f = g$ by assumption (ii), and we have proved that $f = \sum_{k=1}^{\infty} a_k e_k$.

Now assume that (iii) holds. Observe from (2) that we immediately get in the limit as N goes to infinity

$$\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2.$$

Finally, if (iv) holds, then again from (2) we see that $\|f - S_N(f)\|$ converges to 0. Since each $S_N(f)$ is a finite linear combination of elements in $\{e_k\}$, we have completed the circle of implications, and the theorem is proved.

In particular, a closer look at the proof shows that Bessel's inequality holds for any orthonormal family $\{e_k\}$. In contrast, the identity

$$\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2, \quad \text{where } a_k = (f, e_k),$$

which is called **Parseval's identity**, holds if and only if $\{e_k\}_{k=1}^{\infty}$ is also an orthonormal basis.

Now we turn our attention to the existence of a basis.

Theorem 2.4 *Any Hilbert space has an orthonormal basis.*

The first step in the proof of this fact is to recall that (by definition) a Hilbert space $\mathcal{H}$ is separable. Hence, we may choose a countable collection of elements $\mathcal{F} = \{h_k\}$ in $\mathcal{H}$ so that finite linear combinations of elements in $\mathcal{F}$ are dense in $\mathcal{H}$.

We start by recalling a definition already used in the case of finite-dimensional vector spaces. Finitely many elements $g_1, \ldots, g_N$ are said to be **linearly independent** if whenever

$$a_1 g_1 + \cdots + a_N g_N = 0 \quad \text{for some complex numbers } a_i,$$

then $a_1 = a_2 = \cdots = a_N = 0$. In other words, no element g_j is a linear combination of the others. In particular, we note that none of the g_j can be 0. We say that a countable family of elements is **linearly independent** if all finite subsets of this family are linearly independent.

If we next successively disregard the elements h_k that are linearly dependent on the previous elements $h_1, h_2, \ldots, h_{k-1}$, then the resulting collection $h_1 = f_1, f_2, \ldots, f_k, \ldots$ consists of linearly independent elements, whose finite linear combinations are the same as those given by $h_1, h_2, \ldots, h_k, \ldots$, and hence these linear combinations are also dense in $\mathcal{H}$.

The proof of the theorem now follows from an application of a familiar construction called the **Gram-Schmidt process**. Given a finite family of elements $\{f_1, \ldots, f_k\}$ we call the **span** of this family the set of all elements which are finite linear combinations of the elements $\{f_1, \ldots, f_k\}$. We denote the span of $\{f_1, \ldots, f_k\}$ by $\text{Span}(\{f_1, \ldots, f_k\})$.

We now construct a sequence of orthonormal vectors $e_1, e_2, \ldots$ such that $\text{Span}(\{e_1, \ldots, e_n\}) = \text{Span}(\{f_1, \ldots, f_n\})$ for all $n \geq 1$. We do this by induction.

By the linear independence hypothesis, $f_1 \neq 0$, so we may take $e_1 = f_1/\|f_1\|$. Next, assume that orthonormal vectors $e_1, \ldots, e_k$ have been found such that $\text{Span}(\{e_1, \ldots, e_k\}) = \text{Span}(\{f_1, \ldots, f_k\})$ for a given k. We then try e'_{k+1} as $f_{k+1} + \sum_{j=1}^{k} a_j e_j$. To have $(e'_{k+1}, e_j) = 0$ requires that $a_j = -(f_{k+1}, e_j)$, and this choice of a_j for $1 \leq j \leq k$ assures that e'_{k+1} is orthogonal to $e_1, \ldots, e_k$. Moreover our linear independence hypothesis assures that $e'_{k+1} \neq 0$; hence we need only "renormalize" and

take $e_{k+1} = e'_{k+1}/\|e'_{k+1}\|$ to complete the inductive step. With this we have found an orthonormal basis for $\mathcal{H}$

Note that we have implicitly assumed that the number of linearly independent elements $f_1, f_2, \ldots$ is infinite. In the case where there are only N linearly independent vectors $f_1, \ldots, f_N$, then $e_1, \ldots, e_N$ constructed in the same way also provide an orthonormal basis for $\mathcal{H}$. These two cases are differentiated in the following definition. If $\mathcal{H}$ is a Hilbert space with an orthonormal basis consisting of finitely many elements, then we say that $\mathcal{H}$ is **finite-dimensional**. Otherwise $\mathcal{H}$ is said to be **infinite-dimensional**.

2.2 Unitary mappings

A correspondence between two Hilbert spaces that preserves their structure is a unitary transformation. More precisely, suppose we are given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ with respective inner products $(\cdot,\cdot)_{\mathcal{H}}$ and $(\cdot,\cdot)_{\mathcal{H}'}$, and the corresponding norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}'}$. A mapping $U : \mathcal{H} \to \mathcal{H}'$ between these space is called **unitary** if:

(i) U is linear, that is, $U(\alpha f + \beta g) = \alpha U(f) + \beta U(g)$.

(ii) U is a bijection.

(iii) $\|Uf\|_{\mathcal{H}'} = \|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$.

Some observations are in order. First, since U is bijective it must have an inverse $U^{-1} : \mathcal{H}' \to \mathcal{H}$ that is also unitary. Part (iii) above also implies that if U is unitary, then

$$(Uf, Ug)_{\mathcal{H}'} = (f, g)_{\mathcal{H}} \qquad \text{for all } f, g \in \mathcal{H}.$$

To see this, it suffices to "polarize," that is, to note that for any vector space (say over $\mathbb{C}$) with inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$, we have

$$(F, G) = \frac{1}{4}\left[\|F+G\|^2 - \|F-G\|^2 + i\left(\|\frac{F}{i} + G\|^2 - \|\frac{F}{i} - G\|^2\right)\right]$$

whenever F and G are elements of the space.

The above leads us to say that the two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ are **unitarily equivalent** or **unitarily isomorphic** if there exists a unitary mapping $U : \mathcal{H} \to \mathcal{H}'$. Clearly, unitary isomorphism of Hilbert spaces is an equivalence relation.

With this definition we are now in a position to give precise meaning to the statement we made earlier that all infinite-dimensional Hilbert spaces are the same and in that sense $\ell^2(\mathbb{Z})$ in disguise.

Corollary 2.5 *Any two infinite-dimensional Hilbert spaces are unitarily equivalent.*

Proof. If $\mathcal{H}$ and $\mathcal{H}'$ are two infinite-dimensional Hilbert spaces, we may select for each an orthonormal basis, say

$$\{e_1, e_2, \ldots\} \subset \mathcal{H} \quad \text{and} \quad \{e_1', e_2', \ldots\} \subset \mathcal{H}'.$$

Then, consider the mapping defined as follows: if $f = \sum_{k=1}^{\infty} a_k e_k$, then

$$U(f) = g, \quad \text{where} \quad g = \sum_{k=1}^{\infty} a_k e_k'.$$

Clearly, the mapping U is both linear and invertible. Moreover, by Parseval's identity, we must have

$$\|Uf\|_{\mathcal{H}'}^2 = \|g\|_{\mathcal{H}'}^2 = \sum_{k=1}^{\infty} |a_k|^2 = \|f\|_{\mathcal{H}}^2,$$

and the corollary is proved.

Consequently, all infinite-dimensional Hilbert spaces are unitarily equivalent to $\ell^2(\mathbb{N})$, and thus, by relabeling, to $\ell^2(\mathbb{Z})$. By similar reasoning we also have the following:

Corollary 2.6 *Any two finite-dimensional Hilbert spaces are unitarily equivalent if and only if they have the same dimension.*

Thus every finite-dimensional Hilbert space over $\mathbb{C}$ (or over $\mathbb{R}$) is equivalent with $\mathbb{C}^d$ (or $\mathbb{R}^d$), for some d.

2.3 Pre-Hilbert spaces

Although Hilbert spaces arise naturally, one often starts with a **pre-Hilbert space** instead, that is, a space $\mathcal{H}_0$ that satisfies all the defining properties of a Hilbert space except (v); in other words $\mathcal{H}_0$ is not assumed to be complete. A prime example arose implicitly early in the study of Fourier series with the space $\mathcal{H}_0 = \mathcal{R}$ of Riemann integrable functions on $[-\pi, \pi]$ with the usual inner product; we return to this below. Other examples appear in the next chapter in the study of the solutions of partial differential equations.

Fortunately, every pre-Hilbert space $\mathcal{H}_0$ can be completed.

Proposition 2.7 *Suppose we are given a pre-Hilbert space $\mathcal{H}_0$ with inner product $(\cdot,\cdot)_0$. Then we can find a Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)$ such that*

(i) $\mathcal{H}_0 \subset \mathcal{H}$.

(ii) $(f,g)_0 = (f,g)$ *whenever* $f, g \in \mathcal{H}_0$.

(iii) $\mathcal{H}_0$ *is dense in* $\mathcal{H}$.

A Hilbert space satisfying properties like $\mathcal{H}$ in the above proposition is called a **completion** of $\mathcal{H}_0$. We shall only sketch the construction of $\mathcal{H}$, since it follows closely Cantor's familiar method of obtaining the real numbers as the completion of the rationals in terms of Cauchy sequences of rationals.

Indeed, consider the collection of all Cauchy sequences $\{f_n\}$ with $f_n \in \mathcal{H}_0$, $1 \le n < \infty$. One defines an equivalence relation in this collection by saying that $\{f_n\}$ is equivalent to $\{f'_n\}$ if $f_n - f'_n$ converges to 0 as $n \to \infty$. The collection of equivalence classes is then taken to be $\mathcal{H}$. One then easily verifies that $\mathcal{H}$ inherits the structure of a vector space, with an inner product (f,g) defined as $\lim_{n\to\infty}(f_n, g_n)$, where $\{f_n\}$ and $\{g_n\}$ are Cauchy sequences in $\mathcal{H}_0$, representing, respectively, the elements f and g in $\mathcal{H}$. Next, if $f \in \mathcal{H}_0$ we take the sequence $\{f_n\}$, with $f_n = f$ for all n, to represent f as an element of $\mathcal{H}$, giving $\mathcal{H}_0 \subset \mathcal{H}$. To see that $\mathcal{H}$ is complete, let $\{F^k\}_{k=1}^{\infty}$ be a Cauchy sequence in $\mathcal{H}$, with each F^k represented by $\{f_n^k\}_{n=1}^{\infty}$, $f_n^k \in \mathcal{H}_0$. If we define $F \in \mathcal{H}$ as represented by the sequence $\{f_n\}$ with $f_n = f_{N(n)}^n$, where $N(n)$ is so that $|f_{N(n)}^n - f_j^n| \le 1/n$ for $j \ge N(n)$, then we note that $F^k \to F$ in $\mathcal{H}$.

One can also observe that the completion $\mathcal{H}$ of $\mathcal{H}_0$ is unique up to isomorphism. (See Exercise 14.)

3 Fourier series and Fatou's theorem

We have already seen an interesting relation between Hilbert spaces and some elementary facts about Fourier series. Here we want to pursue this idea and also connect it with complex analysis.

When considering Fourier series, it is natural to begin by turning to the broader class of all integrable functions on $[-\pi,\pi]$. Indeed, note that $L^2([-\pi,\pi]) \subset L^1([-\pi,\pi])$, by the Cauchy-Schwarz inequality, since the interval $[-\pi,\pi]$ has finite measure. Thus, if $f \in L^1([-\pi,\pi])$ and $n \in \mathbb{Z}$, we define the nth **Fourier coefficient** of f by

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx.$$

The **Fourier series** of f is then formally $\sum_{n=-\infty}^{\infty} a_n e^{inx}$, and we write

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}$$

to indicate that the sum on the right is the Fourier series of the function on the left. The theory developed thus far provides the natural generalization of some earlier results obtained in Book I.

Theorem 3.1 *Suppose f is integrable on $[-\pi, \pi]$.*

(i) *If $a_n = 0$ for all n, then $f(x) = 0$ for a.e. x.*

(ii) *$\sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{inx}$ tends to $f(x)$ for a.e. x, as $r \to 1$, $r < 1$.*

The second conclusion is the almost everywhere "Abel summability" to f of its Fourier series. Note that since $|a_n| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|\, dx$, the series $\sum a_n r^{|n|} e^{inx}$ converges absolutely and uniformly for each r, $0 \leq r < 1$.

Proof. The first conclusion is an immediate consequence of the second. To prove the latter we recall the identity

$$\sum_{n=-\infty}^{\infty} r^{|n|} e^{iny} = P_r(y) = \frac{1 - r^2}{1 - 2r \cos y + r^2}$$

for the Poisson kernel; see Book I, Chapter 2. Starting with our given $f \in L^1([-\pi, \pi])$ we extend it as a function on $\mathbb{R}$ by making it periodic of period 2π.[3] We then claim that for every x

$$\sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - y) P_r(y)\, dy. \tag{3}$$

Indeed, by the dominated convergence theorem the right-hand side equals

$$\sum r^{|n|} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - y) e^{iny}\, dy.$$

Moreover, for each x and n

$$\begin{aligned}\int_{-\pi}^{\pi} f(x - y) e^{iny}\, dy &= \int_{-\pi+x}^{\pi+x} f(y) e^{in(x-y)}\, dy \\ &= e^{inx} \int_{-\pi}^{\pi} f(y) e^{-iny}\, dy = e^{inx} 2\pi a_n.\end{aligned}$$

[3] Note that we may without loss of generality assume that $f(\pi) = f(-\pi)$ so as to make the periodic extension unambiguous.

The first equality follows by translation invariance (see Section 3, Chapter 2), and the second since $\int_{-\pi}^{\pi} F(y)\,dy = \int_I F(y)\,dy$ whenever F is periodic of period 2π and I is an interval of length 2π (Exercise 3, Chapter 2). With these observations, the identity (3) is established. We can now invoke the facts about approximations to the identity (Theorem 2.1 and Example 4, Chapter 3) to conclude that the left-hand side of (3) tends to $f(x)$ at every point of the Lebesgue set of f, hence almost everywhere. (To be correct, the hypotheses of the theorem require that f be integrable on all of $\mathbb{R}$. We can achieve this for our periodic function by setting f equal to zero outside $[-2\pi, 2\pi]$, and then (3) still holds for this modified f, whenever $x \in [-\pi, \pi]$.)

We return to the more restrictive setting of L^2. We express the essential conclusions of Theorem 2.3 in the context of Fourier series. With $f \in L^2([-\pi, \pi])$, we write as before $a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$.

Theorem 3.2 *Suppose $f \in L^2([-\pi, \pi])$. Then:*

(i) *We have Parseval's relation*

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx.$$

(ii) *The mapping $f \mapsto \{a_n\}$ is a unitary correspondence between $L^2([-\pi, \pi])$ and $\ell^2(\mathbb{Z})$.*

(iii) *The Fourier series of f converges to f in the L^2-norm, that is,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - S_N(f)(x)|^2\,dx \to 0 \qquad \text{as } N \to \infty,$$

where $S_N(f) = \sum_{|n|\le N} a_n e^{inx}$.

To apply the previous results, we let $\mathcal{H} = L^2([-\pi, \pi])$ with inner product $(f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$, and take the orthonormal set $\{e_k\}_{k=1}^{\infty}$ to be the exponentials $\{e^{inx}\}_{n=-\infty}^{\infty}$, with $k = 1$ when $n = 0$, $k = 2n$ for $n > 0$, and $k = 2|n| - 1$ for $n < 0$.

By the previous result, assertion (ii) of Theorem 2.3 holds and thus all the other conclusions hold. We therefore have Parseval's relation, and from (iv) we conclude that $\|f - S_N(f)\|^2 = \sum_{|n|>N} |a_n|^2 \to 0$ as $N \to \infty$. Similarly, if $\{a_n\} \in \ell^2(\mathbb{Z})$ is given, then $\|S_N(f) - S_M(f)\|^2 \to 0$, as $N, M \to \infty$. Hence the completeness of L^2 guarantees that there is an $f \in L^2$ such that $\|f - S_N(f)\| \to 0$, and one verifies directly that f

has $\{a_n\}$ as its Fourier coefficients. Thus we deduce that the mapping $f \mapsto \{a_n\}$ is onto and hence unitary. This is a key conclusion that holds in the setting on L^2 and was not valid in an earlier context of Riemann integrable functions. In fact the space $\mathcal{R}$ of such functions on $[-\pi, \pi]$ is not complete in the norm, containing as it does the continuous functions, but $\mathcal{R}$ is itself restricted to bounded functions.

3.1 Fatou's theorem

Fatou's theorem is a remarkable result in complex analysis. Its proof combines elements of Hilbert spaces, Fourier series, and deeper ideas of differentiation theory, and yet none of these notions appear in its statement. The question that Fatou's theorem answers may be put simply as follows.

> Suppose $F(z)$ is holomorphic in the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. What are conditions on F that guarantee that $F(z)$ will converge, in an appropriate sense, to boundary values $F(e^{i\theta})$ on the unit circle?

In general a holomorphic function in the unit disc can behave quite erratically near the boundary. It turns out, however, that imposing a simple boundedness condition is enough to obtain a strong conclusion.

If F is a function defined in the unit disc $\mathbb{D}$, we say that F has a **radial limit** at the point $-\pi \leq \theta \leq \pi$ on the circle, if the limit

$$\lim_{\substack{r \to 1 \\ r < 1}} F(re^{i\theta})$$

exists.

Theorem 3.3 *A bounded holomorphic function $F(re^{i\theta})$ on the unit disc has radial limits at almost every θ.*

Proof. We know that $F(z)$ has a power series expansion $\sum_{n=0}^{\infty} a_n z^n$ in $\mathbb{D}$ that converges absolutely and uniformly whenever $z = re^{i\theta}$ and $r < 1$. In fact, for $r < 1$ the series $\sum_{n=0}^{\infty} a_n r^n e^{in\theta}$ is the Fourier series of the function $F(re^{i\theta})$, that is,

$$a_n r^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(re^{i\theta}) e^{-in\theta}\, d\theta \qquad \text{when } n \geq 0,$$

and the integral vanishes when $n < 0$. (See also Chapter 3, Section 7 in Book II).

We pick M so that $|F(z)| \leq M$, for all $z \in \mathbb{D}$. By Parseval's identity

$$\sum_{n=0}^{\infty} |a_n|^2 r^{2n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(re^{i\theta})|^2 d\theta \qquad \text{for each } 0 \leq r < 1.$$

Letting $r \to 1$ one sees that $\sum |a_n|^2$ converges (and is $\leq M^2$). We now let $F(e^{i\theta})$ be the L^2-function whose Fourier coefficients are a_n when $n \geq 0$, and 0 when $n < 0$. Hence by conclusion (ii) in Theorem 3.1

$$\sum_{n=0}^{\infty} a_n r^n e^{in\theta} \to F(e^{i\theta}), \qquad \text{for a.e } \theta,$$

concluding the proof of the theorem.

If we examine the argument given above we see that the same conclusion holds for a larger class of functions. In this connection, we define the **Hardy space** $H^2(\mathbb{D})$ to consist of all holomorphic functions F on the unit disc $\mathbb{D}$ that satisfy

$$\sup_{0 \leq r < 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(re^{i\theta})|^2 \, d\theta < \infty.$$

We also define the "norm" for functions F in this class, $\|F\|_{H^2(\mathbb{D})}$, to be the square root of the above quantity.

One notes that if F is bounded, then $F \in H^2(\mathbb{D})$, and moreover the conclusion of the existence of radial limits almost everywhere holds for any $F \in H^2(\mathbb{D})$, by the same argument given for the bounded case.[4] Finally, one notes that $F \in H^2(\mathbb{D})$ if and only if $F(z) = \sum_{n=0}^{\infty} a_n z^n$ with $\sum_{n=0}^{\infty} |a_n|^2 < \infty$; moreover, $\sum_{n=0}^{\infty} |a_n|^2 = \|F\|^2_{H^2(\mathbb{D})}$. This states in particular that $H^2(\mathbb{D})$ is in fact a Hilbert space that can be viewed as the "subspace" $\ell^2(\mathbb{Z}^+)$ of $\ell^2(\mathbb{Z})$, consisting of all $\{a_n\} \in \ell^2(\mathbb{Z})$, with $a_n = 0$ when $n < 0$.

Some general considerations of subspaces and their concomitant orthogonal projections will be taken up next.

4 Closed subspaces and orthogonal projections

A **linear subspace** $\mathcal{S}$ (or simply subspace) of $\mathcal{H}$ is a subset of $\mathcal{H}$ that satisfies $\alpha f + \beta g \in \mathcal{S}$ whenever $f, g \in \mathcal{S}$ and α, β are scalars. In other words, $\mathcal{S}$ is also a vector space. For example in $\mathbb{R}^3$, lines passing through

[4] An even more general statement is given in Problem 5*.

the origin and planes passing through the origin are the one-dimensional and two-dimensional subspaces, respectively.

The subspace $\mathcal{S}$ is **closed** if whenever $\{f_n\} \subset \mathcal{S}$ converges to some $f \in \mathcal{H}$, then f also belongs to $\mathcal{S}$. In the case of finite-dimensional Hilbert spaces, every subspace is closed. This is, however, not true in the general case of infinite-dimensional Hilbert spaces. For instance, as we have already indicated, the subspace of Riemann integrable functions in $L^2([-\pi,\pi])$ is not closed, nor is the subspace obtained by fixing a basis and taking all vectors that are finite linear combinations of these basis elements. It is useful to note that every closed subspace $\mathcal{S}$ of $\mathcal{H}$ is itself a Hilbert space, with the inner product on $\mathcal{S}$ that which is inherited from $\mathcal{H}$. (For the separability of $\mathcal{S}$, see Exercise 11.)

Next, we show that a closed subspace enjoys an important characteristic property of Euclidean geometry.

Lemma 4.1 *Suppose $\mathcal{S}$ is a closed subspace of $\mathcal{H}$ and $f \in \mathcal{H}$. Then:*

(i) *There exists a (unique) element $g_0 \in \mathcal{S}$ which is closest to f, in the sense that*

$$\|f - g_0\| = \inf_{g\in\mathcal{S}} \|f - g\|.$$

(ii) *The element $f - g_0$ is perpendicular to $\mathcal{S}$, that is,*

$$(f - g_0, g) = 0 \quad \textit{for all } g \in \mathcal{S}.$$

The situation in the lemma can be visualized as in Figure 1.

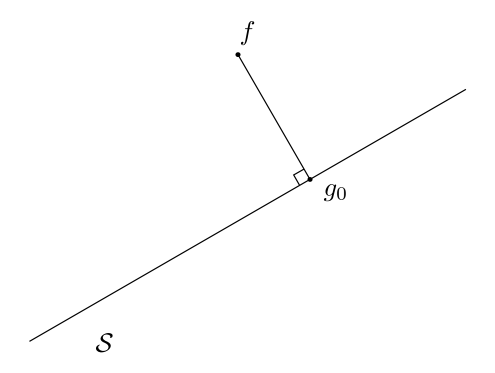


Figure 1. Nearest element to f in $\mathcal{S}$

Proof. If $f \in \mathcal{S}$, then we choose $f = g_0$, and there is nothing left to prove. Otherwise, we let $d = \inf_{g \in \mathcal{S}} \|f - g\|$, and note that we must have $d > 0$ since $f \notin \mathcal{S}$ and $\mathcal{S}$ is closed. Consider a sequence $\{g_n\}_{n=1}^{\infty}$ in $\mathcal{S}$ such that

$$\|f - g_n\| \to d \qquad \text{as } n \to \infty.$$

We claim that $\{g_n\}$ is a Cauchy sequence whose limit will be the desired element g_0. In fact, it would suffice to show that a subsequence of $\{g_n\}$ converges, and this is immediate in the finite-dimensional case because a closed ball is compact. However, in general this compactness fails, as we shall see in Section 6, and so a more intricate argument is needed at this point.

To prove our claim, we use the **parallelogram law**, which states that in a Hilbert space $\mathcal{H}$

$$\|A + B\|^2 + \|A - B\|^2 = 2\left[\|A\|^2 + \|B\|^2\right] \qquad \text{for all } A, B \in \mathcal{H}. \tag{4}$$

The simple verification of this equality, which consists of writing each norm in terms of the inner product, is left to the reader. Putting $A = f - g_n$ and $B = f - g_m$ in the parallelogram law, we find

$$\|2f - (g_n + g_m)\|^2 + \|g_m - g_n\|^2 = 2\left[\|f - g_n\|^2 + \|f - g_m\|^2\right].$$

However $\mathcal{S}$ is a subspace, so the quantity $\frac{1}{2}(g_n + g_m)$ belongs to $\mathcal{S}$, hence

$$\|2f - (g_n + g_m)\| = 2\|f - \frac{1}{2}(g_n + g_m)\| \geq 2d.$$

Therefore

$$\begin{aligned}\|g_m - g_n\|^2 &= 2\left[\|f - g_n\|^2 + \|f - g_m\|^2\right] - \|2f - (g_n + g_m)\|^2 \\ &\leq 2\left[\|f - g_n\|^2 + \|f - g_m\|^2\right] - 4d^2.\end{aligned}$$

By construction, we know that $\|f - g_n\| \to d$ and $\|f - g_m\| \to d$ as $n, m \to \infty$, so the above inequality implies that $\{g_n\}$ is a Cauchy sequence. Since $\mathcal{H}$ is complete and $\mathcal{S}$ closed, the sequence $\{g_n\}$ must have a limit g_0 in $\mathcal{S}$, and then it satisfies $d = \|f - g_0\|$.

We prove that if $g \in \mathcal{S}$, then $g \perp (f - g_0)$. For each ϵ (positive or negative), consider the perturbation of g_0 defined by $g_0 - \epsilon g$. This element belongs to $\mathcal{S}$, hence

$$\|f - (g_0 - \epsilon g)\|^2 \geq \|f - g_0\|^2.$$

Since $\|f - (g_0 - \epsilon g)\|^2 = \|f - g_0\|^2 + \epsilon^2\|g\|^2 + 2\epsilon\,\mathrm{Re}(f - g_0, g)$, we find that

$$2\epsilon\,\mathrm{Re}(f - g_0, g) + \epsilon^2\|g\|^2 \geq 0. \tag{5}$$

If $\mathrm{Re}(f - g_0, g) < 0$, then taking ϵ small and positive contradicts (5). If $\mathrm{Re}(f - g_0, g) > 0$, a contradiction also follows by taking ϵ small and negative. Thus $\mathrm{Re}(f - g_0, g) = 0$. By considering the perturbation $g_0 - i\epsilon g$, a similar argument gives $\mathrm{Im}(f - g_0, g) = 0$, and hence $(f - g_0, g) = 0$.

Finally, the uniqueness of g_0 follows from the above observation about orthogonality. Suppose $\tilde{g}_0$ is another point in $\mathcal{S}$ that minimizes the distance to f. By taking $g = g_0 - \tilde{g}_0$ in our last argument we find $(f - g_0) \perp (g_0 - \tilde{g}_0)$, and the Pythagorean theorem gives

$$\|f - \tilde{g}_0\|^2 = \|f - g_0\|^2 + \|g_0 - \tilde{g}_0\|^2.$$

Since by assumption $\|f - \tilde{g}_0\|^2 = \|f - g_0\|^2$, we conclude that $\|g_0 - \tilde{g}_0\| = 0$, as desired.

Using the lemma, we may now introduce a useful concept that is another expression of the notion of orthogonality. If $\mathcal{S}$ is a subspace of a Hilbert space $\mathcal{H}$, we define the **orthogonal complement** of $\mathcal{S}$ by

$$\mathcal{S}^\perp = \{f \in \mathcal{H} : (f, g) = 0 \quad \text{for all } g \in \mathcal{S}\}.$$

Clearly, $S^\perp$ is also a subspace of $\mathcal{H}$, and moreover $\mathcal{S} \cap \mathcal{S}^\perp = \{0\}$. To see this, note that if $f \in \mathcal{S} \cap \mathcal{S}^\perp$, then f must be orthogonal to itself; thus $0 = (f, f) = \|f\|$, and therefore $f = 0$. Moreover, $\mathcal{S}^\perp$ is itself a closed subspace. Indeed, if $f_n \to f$, then $(f_n, g) \to (f, g)$ for every g, by the Cauchy-Schwarz inequality. Hence if $(f_n, g) = 0$ for all $g \in \mathcal{S}$ and all n, then $(f, g) = 0$ for all those g.

Proposition 4.2 *If $\mathcal{S}$ is a closed subspace of a Hilbert space $\mathcal{H}$, then*

$$\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp.$$

The notation in the proposition means that every $f \in \mathcal{H}$ can be written uniquely as $f = g + h$, where $g \in \mathcal{S}$ and $h \in \mathcal{S}^\perp$; we say that $\mathcal{H}$ is the **direct sum** of S and $S^\perp$. This is equivalent to saying that any f in $\mathcal{H}$ is the sum of two elements, one in $\mathcal{S}$, the other in $\mathcal{S}^\perp$, and that $\mathcal{S} \cap \mathcal{S}^\perp$ contains only 0.

The proof of the proposition relies on the previous lemma giving the closest element of f in $\mathcal{S}$. In fact, for any $f \in \mathcal{H}$, we choose g_0 as in the

lemma and write

$$f = g_0 + (f - g_0).$$

By construction $g_0 \in \mathcal{S}$, and the lemma implies $f - g_0 \in S^\perp$, and this shows that f is the sum of an element in $\mathcal{S}$ and one in $\mathcal{S}^\perp$. To prove that this decomposition is unique, suppose that

$$f = g + h = \tilde{g} + \tilde{h} \qquad \text{where } g, \tilde{g} \in \mathcal{S} \text{ and } h, \tilde{h} \in \mathcal{S}^\perp.$$

Then, we must have $g - \tilde{g} = \tilde{h} - h$. Since the left-hand side belongs to $\mathcal{S}$ while the right-hand side belongs to $\mathcal{S}^\perp$ the fact that $\mathcal{S} \cap \mathcal{S}^\perp = \{0\}$ implies $g - \tilde{g} = 0$ and $\tilde{h} - h = 0$. Therefore $g = \tilde{g}$ and $h = \tilde{h}$ and the uniqueness is established.

With the decomposition $\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp$ one has the natural projection onto S defined by

$$P_{\mathcal{S}}(f) = g, \qquad \text{where } f = g + h \text{ and } g \in \mathcal{S},\ h \in \mathcal{S}^\perp.$$

The mapping $P_{\mathcal{S}}$ is called the **orthogonal projection** onto $\mathcal{S}$ and satisfies the following simple properties:

(i) $f \mapsto P_{\mathcal{S}}(f)$ is linear,

(ii) $P_{\mathcal{S}}(f) = f$ whenever $f \in \mathcal{S}$,

(iii) $P_{\mathcal{S}}(f) = 0$ whenever $f \in \mathcal{S}^\perp$,

(iv) $\|P_{\mathcal{S}}(f)\| \leq \|f\|$ for all $f \in \mathcal{H}$.

Property (i) means that $P_{\mathcal{S}}(\alpha f_1 + \beta f_2) = \alpha P_{\mathcal{S}}(f_1) + \beta P_{\mathcal{S}}(f_2)$, whenever $f_1, f_2 \in \mathcal{H}$ and α and β are scalars.

It will be useful to observe the following. Suppose $\{e_k\}$ is a (finite or infinite) collection of orthonormal vectors in $\mathcal{H}$. Then the orthogonal projection P in the closure of the subspace spanned by $\{e_k\}$ is given by $P(f) = \sum_k (f, e_k)e_k$. In case the collection is infinite, the sum converges in the norm of $\mathcal{H}$.

We illustrate this with two examples that arise in Fourier analysis.

Example 1. On $L^2([-\pi, \pi])$, recall that if $f(\theta) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ then the partial sums of the Fourier series are

$$S_N(f)(\theta) = \sum_{n=-N}^{N} a_n e^{in\theta}.$$

Therefore, the partial sum operator S_N consists of the projection onto the closed subspace spanned by $\{e_{-N}, \ldots, e_N\}$.

The sum S_N can be realized as a convolution

$$S_N(f)(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(\theta - \varphi) f(\varphi)\, d\varphi,$$

where $D_N(\theta) = \sin((N + 1/2)\theta)/\sin(\theta/2)$ is the **Dirichlet kernel**.

EXAMPLE 2. Once again, consider $L^2([-\pi, \pi])$ and let $\mathcal{S}$ denote the subspace that consists of all $F \in L^2([-\pi, \pi])$ with

$$F(\theta) \sim \sum_{n=0}^{\infty} a_n e^{in\theta}.$$

In other words, $\mathcal{S}$ is the space of square integrable functions whose Fourier coefficients a_n vanish for $n < 0$. From the proof of Fatou's theorem, this implies that $\mathcal{S}$ can be identified with the Hardy space $H^2(\mathbb{D})$, where $\mathbb{D}$ is the unit disc, and so is a closed subspace unitarily isomorphic to $\ell^2(\mathbb{Z}^+)$. Therefore, using this identification, if P denotes the orthogonal projection from $L^2([-\pi, \pi])$ to $\mathcal{S}$, we may also write $P(f)(z)$ for the element corresponding to $H^2(\mathbb{D})$, that is,

$$P(f)(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Given $f \in L^2([-\pi, \pi])$, we define the **Cauchy integral** of f by

$$C(f)(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{\zeta - z}\, d\zeta,$$

where γ denotes the unit circle and z belongs to the unit disc. Then we have the identity

$$P(f)(z) = C(f)(z), \qquad \text{for all } z \in \mathbb{D}.$$

Indeed, since $f \in L^2$ it follows by the Cauchy-Schwarz inequality that $f \in L^1([-\pi, \pi])$, and therefore we may interchange the sum and integral

in the following calculation (recall $|z| < 1$):

$$\begin{aligned} P(f)(z) = \sum_{n=0}^{\infty} a_n z^n &= \sum_{n=0}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta} d\theta \right) z^n \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) \sum_{n=0}^{\infty} (e^{-i\theta} z)^n d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(e^{i\theta})}{1 - e^{-i\theta} z} d\theta \\ &= \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{f(e^{i\theta})}{e^{i\theta} - z} i e^{i\theta} d\theta \\ &= C(f)(z). \end{aligned}$$

5 Linear transformations

The focus of analysis in Hilbert spaces is largely the study of their linear transformations. We have already encountered two classes of such transformations, the unitary mappings and the orthogonal projections. There are two other important classes we shall deal with in this chapter in some detail: the "linear functionals" and the "compact operators," and in particular those that are symmetric.

Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are two Hilbert spaces. A mapping $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a **linear transformation** (also called **linear operator** or **operator**) if

$$T(af + bg) = aT(f) + bT(g) \quad \text{for all scalars } a, b \text{ and } f, g \in \mathcal{H}_1.$$

Clearly, linear operators satisfy $T(0) = 0$.

We shall say that a linear operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is **bounded** if there exists $M > 0$ so that

$$\|T(f)\|_{\mathcal{H}_2} \leq M \|f\|_{\mathcal{H}_1}. \tag{6}$$

The **norm** of T is denoted by $\|T\|_{\mathcal{H}_1 \to \mathcal{H}_2}$ or simply $\|T\|$ and defined by

$$\|T\| = \inf M,$$

where the infimum is taken over all M so that (6) holds. A trivial example is given by the **identity operator** I, with $I(f) = f$. It is of course a unitary operator and a projection, with $\|I\| = 1$.

In what follows we shall generally drop the subscripts attached to the norms of elements of a Hilbert space, when this causes no confusion.

Lemma 5.1 $\|T\| = \sup\{|(Tf,g)| : \ \|f\| \leq 1, \ \|g\| \leq 1\}$, *where of course* $f \in \mathcal{H}_1$ *and* $g \in \mathcal{H}_2$.

Proof. If $\|T\| \leq M$, the Cauchy-Schwarz inequality gives

$$|(Tf,g)| \leq M \qquad \text{whenever } \|f\| \leq 1 \text{ and } \|g\| \leq 1;$$

thus $\sup\{|(Tf,g)| : \ \|f\| \leq 1, \ \|g\| \leq 1\} \leq \|T\|$.

Conversely, if $\sup\{|(Tf,g)| : \ \|f\| \leq 1, \ \|g\| \leq 1\} \leq M$, we claim that $\|Tf\| \leq M\|f\|$ for all f. If f or Tf is zero, there is nothing to prove. Otherwise, $f' = f/\|f\|$ and $g' = Tf/\|Tf\|$ have norm 1, so by assumption

$$|(Tf',g')| \leq M.$$

But since $|(Tf',g')| = \|Tf\|/\|f\|$ this gives $\|Tf\| \leq M\|f\|$, and the lemma is proved.

A linear transformation T is **continuous** if $T(f_n) \to T(f)$ whenever $f_n \to f$. Clearly, linearity implies that T is continuous on all of $\mathcal{H}_1$ if and only if it is continuous at the origin. In fact, the conditions of being bounded or continuous are equivalent.

Proposition 5.2 *A linear operator* $T : \mathcal{H}_1 \to \mathcal{H}_2$ *is bounded if and only if it is continuous.*

Proof. If T is bounded, then $\|T(f) - T(f_n)\|_{\mathcal{H}_2} \leq M\|f - f_n\|_{\mathcal{H}_1}$, hence T is continuous. Conversely, suppose that T is continuous but not bounded. Then for each n there exists $f_n \neq 0$ such that $\|T(f_n)\| \geq n\|f_n\|$. The element $g_n = f_n/(n\|f_n\|)$ has norm $1/n$, hence $g_n \to 0$. Since T is continuous at 0, we must have $T(g_n) \to 0$, which contradicts the fact that $\|T(g_n)\| \geq 1$. This proves the proposition.

In the rest of this chapter we shall assume that all linear operators are bounded, hence continuous. It is noteworthy to recall that any linear operator between finite-dimensional Hilbert spaces is necessarily continuous.

5.1 Linear functionals and the Riesz representation theorem

A **linear functional** ℓ is a linear transformation from a Hilbert space $\mathcal{H}$ to the underlying field of scalars, which we may assume to be the

complex numbers,

$$\ell : \mathcal{H} \to \mathbb{C}.$$

Of course, we view $\mathbb{C}$ as a Hilbert space equipped with its standard norm, the absolute value.

A natural example of a linear functional is provided by the inner product on $\mathcal{H}$. Indeed, for fixed $g \in \mathcal{H}$, the map

$$\ell(f) = (f, g)$$

is linear, and also bounded by the Cauchy-Schwarz inequality. Indeed,

$$|(f, g)| \leq M\|f\|, \qquad \text{where } M = \|g\|.$$

Moreover, $\ell(g) = M\|g\|$ so we have $\|\ell\| = \|g\|$. The remarkable fact is that this example is exhaustive, in the sense that every *continuous* linear functional on a Hilbert space arises as an inner product. This is the so-called Riesz representation theorem.

Theorem 5.3 *Let ℓ be a continuous linear functional on a Hilbert space $\mathcal{H}$. Then, there exists a unique $g \in \mathcal{H}$ such that*

$$\ell(f) = (f, g) \qquad \text{for all } f \in \mathcal{H}.$$

Moreover, $\|\ell\| = \|g\|$.

Proof. Consider the subspace of $\mathcal{H}$ defined by

$$\mathcal{S} = \{f \in \mathcal{H} : \ell(f) = 0\}.$$

Since ℓ is continuous the subspace $\mathcal{S}$, which is called the **null-space** of ℓ, is closed. If $\mathcal{S} = \mathcal{H}$, then $\ell = 0$ and we take $g = 0$. Otherwise $\mathcal{S}^\perp$ is non-trivial and we may pick any $h \in \mathcal{S}^\perp$ with $\|h\| = 1$. With this choice of h we determine g by setting $g = \overline{\ell(h)}h$. Thus if we let $u = \ell(f)h - \ell(h)f$, then $u \in \mathcal{S}$, and therefore $(u, h) = 0$. Hence

$$0 = (\ell(f)h - \ell(h)f, h) = \ell(f)(h, h) - (f, \overline{\ell(h)}h).$$

Since $(h, h) = 1$, we find that $\ell(f) = (f, g)$ as desired.

At this stage we record the following remark for later use. Let $\mathcal{H}_0$ be a pre-Hilbert space whose completion is $\mathcal{H}$. Suppose ℓ_0 is a linear functional on $\mathcal{H}_0$ which is bounded, that is, $|\ell_0(f)| \leq M\|f\|$ for all $f \in$

$\mathcal{H}_0$. Then ℓ_0 has an extension ℓ to a bounded linear functional on $\mathcal{H}$, with $|\ell(f)| \leq M\|f\|$ for all $f \in \mathcal{H}$. This extension is also unique. To see this, one merely notes that $\{\ell_0(f_n)\}$ is a Cauchy sequence whenever the vectors $\{f_n\}$ belong to $\mathcal{H}_0$, and $f_n \to f$ in $\mathcal{H}$, as $n \to \infty$. Thus we may define $\ell(f)$ as $\lim_{n\to\infty} \ell_0(f_n)$. The verification of the asserted properties of ℓ is then immediate. (This result is a special case of the extension Lemma 1.3 in the next chapter.)

5.2 Adjoints

The first application of the Riesz representation theorem is to determine the existence of the "adjoint" of a linear transformation.

Proposition 5.4 *Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear transformation. There exists a unique bounded linear transformation T^* on $\mathcal{H}$ so that:*

(i) $(Tf, g) = (f, T^*g)$,

(ii) $\|T\| = \|T^*\|$,

(iii) $(T^*)^* = T$.

The linear operator $T^* : \mathcal{H} \to \mathcal{H}$ satisfying the above conditions is called the **adjoint** of T.

To prove the existence of an operator satisfying (i) above, we observe that for each fixed $g \in \mathcal{H}$, the linear functional $\ell = \ell_g$, defined by

$$\ell(f) = (Tf, g),$$

is bounded. Indeed, since T is bounded one has $\|Tf\| \leq M\|f\|$; hence the Cauchy-Schwarz inequality implies that

$$|\ell(f)| \leq \|Tf\|\,\|g\| \leq B\|f\|,$$

where $B = M\|g\|$. Consequently, the Riesz representation theorem guarantees the existence of a unique $h \in \mathcal{H}$, $h = h_g$, such that

$$\ell(f) = (f, h).$$

Then we define $T^*g = h$, and note that the association $T^* : g \mapsto h$ is linear and satisfies (i).

The fact that $\|T\| = \|T^*\|$ follows at once from (i) and Lemma 5.1:

$$\begin{aligned}\|T\| &= \sup\{|(Tf, g)| :\ \|f\| \leq 1,\ \|g\| \leq 1\}\\ &= \sup\{|(f, T^*g)| :\ \|f\| \leq 1,\ \|g\| \leq 1\}\\ &= \|T^*\|.\end{aligned}$$

To prove (iii), note that $(Tf, g) = (f, T^*g)$ for all f and g if and only if $(T^*f, g) = (f, Tg)$ for all f and g, as one can see by taking complex conjugates and reversing the roles of f and g.

We record here a few additional remarks.

(a) In the special case when $T = T^*$ (we say that T is **symmetric**), then

$$\|T\| = \sup\{|(Tf, f)| : \ \|f\| = 1\}. \tag{7}$$

This should be compared to Lemma 5.1, which holds for any linear operator. To establish (7), let $M = \sup\{|(Tf, f)| : \ \|f\| = 1\}$. By Lemma 5.1 it is clear that $M \leq \|T\|$. Conversely, if f and g belong on $\mathcal{H}$, then one has the following "polarization" identity which is easy to verify

$$\begin{aligned}(Tf, g) = &\frac{1}{4}[(T(f + g), f + g) - (T(f - g), f - g) \\ &+ i\,(T(f + ig), f + ig) - i\,(T(f - ig), f - ig)].\end{aligned}$$

For any $h \in \mathcal{H}$, the quantity (Th, h) is real, because $T = T^*$, hence $(Th, h) = (h, T^*h) = (h, Th) = \overline{(Th, h)}$. Consequently

$$\mathrm{Re}(Tf, g) = \frac{1}{4}\left[(T(f + g), f + g) - (T(f - g), f - g)\right].$$

Now $|(Th, h)| \leq M\|h\|^2$, so $|\mathrm{Re}(Tf, g)| \leq \frac{M}{4}\left[\|f + g\|^2 + \|f - g\|^2\right]$, and an application of the parallelogram law (4) then implies

$$|\mathrm{Re}(Tf, g)| \leq \frac{M}{2}[\|f\|^2 + \|g\|^2].$$

So if $\|f\| \leq 1$ and $\|g\| \leq 1$, then $|\mathrm{Re}(Tf, g)| \leq M$. In general, we may replace g by $e^{i\theta}g$ in the last inequality to find that whenever $\|f\| \leq 1$ and $\|g\| \leq 1$, then $|(Tf, g)| \leq M$, and invoking Lemma 5.1 once again gives the result, $\|T\| \leq M$.

(b) Let us note that if T and S are bounded linear transformations of $\mathcal{H}$ to itself, then so is their product TS, defined by $(TS)(f) = T(S(f))$. Moreover we have automatically $(TS)^* = S^*T^*$; in fact, $(TSf, g) = (Sf, T^*g) = (f, S^*T^*g)$.

(c) One can also exhibit a natural connection between linear transformations on a Hilbert space and their associated bilinear forms. Suppose first that T is a bounded operator in $\mathcal{H}$. Define the corresponding bilinear form B by

$$B(f, g) = (Tf, g). \tag{8}$$

Note that B is linear in f and conjugate linear in g. Also by the Cauchy-Schwarz inequality $|B(f,g)| \leq M\|f\|\,\|g\|$, where $M = \|T\|$. Conversely if B is linear in f, conjugate linear in g and satisfies $|B(f,g)| \leq M\|f\|\,\|g\|$, there is a unique linear transformation so that (8) holds with $M = \|T\|$. This can be proved by the argument of Proposition 5.4; the details are left to the reader.

5.3 Examples

Having presented the elementary facts about Hilbert spaces, we now digress to describe briefly the background of some of the early developments of the theory. A motivating problem of considerable interest was that of the study of the "eigenfunction expansion" of a differential operator L. A particular case, that of a Sturm-Liouville operator, arises on an interval $[a,b]$ of $\mathbb{R}$ with L defined by

$$L = \frac{d^2}{dx^2} - q(x),$$

where q is a given real-valued function. The question is then that of expanding an "arbitrary" function in terms of the eigenfunctions φ, that is those functions that satisfy $L(\varphi) = \mu\varphi$ for some $\mu \in \mathbb{R}$. The classical example of this is that of Fourier series, where $L = d^2/dx^2$ on the interval $[-\pi,\pi]$ with each exponential e^{inx} an eigenfunction of L with eigenvalue $\mu = -n^2$.

When made precise in the "regular" case, the problem for L can be resolved by considering an associated "integral operator" T defined on $L^2([a,b])$ by

$$T(f)(x) = \int_a^b K(x,y)f(y)\,dy,$$

with the property that for suitable f,

$$LT(f) = f.$$

It turns out that a key feature that makes the study of T tractable is a certain compactness it enjoys. We now pass to the definitions and elaboration of some of these ideas, and begin by giving two relevant illustrations of classes of operators on Hilbert spaces.

Infinite diagonal matrix

Suppose $\{\varphi_k\}_{k=1}^{\infty}$ is an orthonormal basis of $\mathcal{H}$. Then, a linear transformation $T : \mathcal{H} \to \mathcal{H}$ is said to be **diagonalized** with respect to the basis

$\{\varphi_k\}$ if

$$T(\varphi_k) = \lambda_k \varphi_k, \qquad \text{where } \lambda_k \in \mathbb{C} \text{ for all } k.$$

In general, a *non-zero* element φ is called an **eigenvector** of T with **eigenvalue** λ if $T\varphi = \lambda\varphi$. So the φ_k above are eigenvectors of T, and the numbers λ_k are the corresponding eigenvalues.

So if

$$f \sim \sum_{k=1}^{\infty} a_k \varphi_k \qquad \text{then} \qquad Tf \sim \sum_{k=1}^{\infty} a_k \lambda_k \varphi_k.$$

The sequence $\{\lambda_k\}$ is called the **multiplier sequence** corresponding to T.

In this case, one can easily verify the following facts:

- $\|T\| = \sup_k |\lambda_k|$.
- T^* corresponds to the sequence $\{\overline{\lambda}_k\}$; hence $T = T^*$ if and only if the λ_k are real.
- T is unitary if and only if $|\lambda_k| = 1$ for all k.
- T is an orthogonal projection if and only if $\lambda_k = 0$ or 1 for all k.

As a particular example, consider $\mathcal{H} = L^2([-\pi, \pi])$, and assume that every $f \in L^2([-\pi, \pi])$ is extended to $\mathbb{R}$ by periodicity, so that $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. Let $\varphi_k(x) = e^{ikx}$ for $k \in \mathbb{Z}$. For a fixed $h \in \mathbb{R}$ the operator U_h defined by

$$U_h(f)(x) = f(x + h)$$

is unitary with $\lambda_k = e^{ikh}$. Hence

$$U_h(f) \sim \sum_{k=-\infty}^{\infty} a_k \lambda_k e^{ikx} \qquad \text{if} \qquad f \sim \sum_{k=-\infty}^{\infty} a_k e^{ikx}.$$

Integral operators, and in particular, Hilbert-Schmidt operators

Let $\mathcal{H} = L^2(\mathbb{R}^d)$. If we can define an operator $T : \mathcal{H} \to \mathcal{H}$ by the formula

$$T(f)(x) = \int_{\mathbb{R}^d} K(x, y) f(y)\, dy \qquad \text{whenever } f \in L^2(\mathbb{R}^d),$$

we say that the operator T is an **integral operator** and K is its associated **kernel**.

In fact, it was the problem of invertibility related to such operators, and more precisely the question of solvability of the equation $f - Tf = g$ for given g, that initiated the study of Hilbert spaces. These equations were then called "integral equations."

In general a bounded linear transformation cannot be expressed as an (absolutely convergent) integral operator. However, there is an interesting class for which this is possible and which has a number of other worthwhile properties: **Hilbert-Schmidt operators**, those with a kernel K that belongs to $L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

Proposition 5.5 *Let T be a Hilbert-Schmidt operator on $L^2(\mathbb{R}^d)$ with kernel K.*

(i) *If $f \in L^2(\mathbb{R}^d)$, then for almost every x the function $y \mapsto K(x,y)f(y)$ is integrable.*

(ii) *The operator T is bounded from $L^2(\mathbb{R}^d)$ to itself, and*

$$\|T\| \leq \|K\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)},$$

where $\|K\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)}$ is the L^2-norm of K on $\mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{2d}$.

(iii) *The adjoint T^* has kernel $\overline{K(y,x)}$.*

Proof. By Fubini's theorem we know that for almost every x, the function $y \mapsto |K(x,y)|^2$ is integrable. Then, part (i) follows directly from an application of the Cauchy-Schwarz inequality.

For (ii), we make use again of the Cauchy-Schwarz inequality as follows

$$\begin{aligned}\left|\int K(x,y)f(y)\,dy\right| &\leq \int |K(x,y)||f(y)|\,dy \\ &\leq \left(\int |K(x,y)|^2\,dy\right)^{1/2}\left(\int |f(y)|^2\,dy\right)^{1/2}.\end{aligned}$$

Therefore, squaring this and integrating in x yields

$$\begin{aligned}\|Tf\|^2_{L^2(\mathbb{R}^d)} &\leq \int\left(\int |K(x,y)|^2 dy \int |f(y)|^2 dy\right) dx \\ &= \|K\|^2_{L^2(\mathbb{R}^d \times \mathbb{R}^d)}\|f\|^2_{L^2(\mathbb{R}^d)}.\end{aligned}$$

Finally, part (iii) follows by writing out (Tf, g) in terms of a double integral, and then interchanging the order of integration, as is permissible by Fubini's theorem.

Hilbert-Schmidt operators can be defined analogously for the Hilbert space $L^2(E)$, where E is a measurable subset of $\mathbb{R}^d$. We leave it to the reader to formulate an prove the analogue of Proposition 5.5 that holds in this case.

Hilbert-Schmidt operators enjoy another important property: they are compact. We will now discuss this feature in more detail.

6 Compact operators

We shall use the notion of sequential compactness in a Hilbert space $\mathcal{H}$: a set $X \subset \mathcal{H}$ is **compact** if for every sequence $\{f_n\}$ in X, there exists a subsequence $\{f_{n_k}\}$ that converges in the norm to an element *in* X.

Let $\mathcal{H}$ denote a Hilbert space, and B the closed unit ball in $\mathcal{H}$,

$$B = \{f \in \mathcal{H} : \|f\| \leq 1\}.$$

A well-known result in elementary real analysis says that in a finite-dimensional Euclidean space, a closed and bounded set is compact. However, this does not carry over to the infinite-dimensional case. The fact is that in this case the unit ball, while closed and bounded, is not compact. To see this, consider the sequence $\{f_n\} = \{e_n\}$, where the e_n are orthonormal. By the Pythagorean theorem, $\|e_n - e_m\|^2 = 2$ if $n \neq m$, so no subsequence of the $\{e_n\}$ can converge.

In the infinite-dimensional case we say that a linear operator $T : \mathcal{H} \to \mathcal{H}$ is **compact** if the closure of

$$T(B) = \{g \in \mathcal{H} : g = T(f) \text{ for some } f \in B\}$$

is a compact set. Equivalently, an operator T is compact if, whenever $\{f_k\}$ is a bounded sequence in $\mathcal{H}$, there exists a subsequence $\{f_{n_k}\}$ so that Tf_{n_k} converges. Note that a compact operator is automatically bounded.

Note that by what has been said, a linear transformation is in general not compact (take for instance the identity operator!). However, if T is of **finite rank**, which means that its range is finite-dimensional, then it is automatically compact. It turns out that dealing with compact operators provides us with the closest analogy to the usual theorems of (finite-dimensional) linear algebra. Some relevant analytic properties of compact operators are given by the proposition below.

Proposition 6.1 *Suppose T is a bounded linear operator on $\mathcal{H}$.*

(i) *If S is compact on $\mathcal{H}$, then ST and TS are also compact.*

(ii) *If $\{T_n\}$ is a family of compact linear operators with $\|T_n - T\| \to 0$ as n tends to infinity, then T is compact.*

(iii) *Conversely, if T is compact, there is a sequence $\{T_n\}$ of operators of finite rank such that $\|T_n - T\| \to 0$.*

(iv) *T is compact if and only if T^* is compact.*

Proof. Part (i) is immediate. For part (ii) we use a diagonalization argument. Suppose $\{f_k\}$ is a bounded sequence in $\mathcal{H}$. Since T_1 is compact, we may extract a subsequence $\{f_{1,k}\}_{k=1}^{\infty}$ of $\{f_k\}$ such that $T_1(f_{1,k})$ converges. From $\{f_{1,k}\}$ we may find a subsequence $\{f_{2,k}\}_{k=1}^{\infty}$ such that $T_2(f_{2,k})$ converges, and so on. If we let $g_k = f_{k,k}$, then we claim $\{T(g_k)\}$ is a Cauchy sequence. We have

$$\|T(g_k) - T(g_\ell)\| \leq \|T(g_k) - T_m(g_k)\| + \|T_m(g_k) - T_m(g_\ell)\| + \\ + \|T_m(g_\ell) - T(g_\ell)\|.$$

Since $\|T - T_m\| \to 0$ and $\{g_k\}$ is bounded, we can make the first and last term each $< \epsilon/3$ for some large m independent of k and ℓ. With this fixed m, we note that by construction $\|T_m(g_k) - T_m(g_\ell)\| < \epsilon/3$ for all large k and ℓ. This proves our claim; hence $\{T(g_k)\}$ converges in $\mathcal{H}$.

To prove (iii) let $\{e_k\}_{k=1}^{\infty}$ be a basis of $\mathcal{H}$ and let Q_n be the orthogonal projection on the subspace spanned by the e_k with $k > n$. Then clearly $Q_n(g) \sim \sum_{k>n} a_k e_k$ whenever $g \sim \sum_{k=1}^{\infty} a_k e_k$, and $\|Q_n g\|^2$ is a decreasing sequence that tends to 0 as $n \to \infty$ for any $g \in \mathcal{H}$. We claim that $\|Q_n T\| \to 0$ as $n \to \infty$. If not, there is a $c > 0$ so that $\|Q_n T\| \geq c$, and hence for each n we can find f_n, with $\|f_n\| = 1$ so that $\|Q_n T f_n\| \geq c$. Now by compactness of T, choosing an appropriate subsequence $\{f_{n_k}\}$, we have $Tf_{n_k} \to g$ for some g. But $Q_{n_k}(g) = Q_{n_k} T f_{n_k} - Q_{n_k}(Tf_{n_k} - g)$, and hence we conclude that $\|Q_{n_k}(g)\| \geq c/2$, for large k. This contradiction shows that $\|Q_n T\| \to 0$. So if P_n is the complementary projection on the finite-dimensional space spanned by $e_1, \ldots, e_n$, $I = P_n + Q_n$, then $\|Q_n T\| \to 0$ means that $\|P_n T - T\| \to 0$. Since each $P_n T$ is of finite rank, assertion (iii) is established.

Finally, if T is compact the fact that $\|P_n T - T\| \to 0$ implies $\|T^* P_n - T^*\| \to 0$, and clearly $T^* P_n$ is again of finite rank. Thus we need only appeal to the second conclusion to prove the last.

We now state two further observations about compact operators.

- If T can be diagonalized with respect to some basis $\{\varphi_k\}$ of eigenvectors and corresponding eigenvalues $\{\lambda_k\}$, then T is compact if and only if $|\lambda_k| \to 0$. See Exercise 25.

- Every Hilbert-Schmidt operator is compact.

To prove the second point, recall that a Hilbert-Schmidt operator is given on $L^2(\mathbb{R}^d)$ by

$$T(f)(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\, dy, \qquad \text{where } K \in L^2(\mathbb{R}^d \times \mathbb{R}^d).$$

If $\{\varphi_k\}_{k=1}^\infty$ denotes an orthonormal basis for $L^2(\mathbb{R}^d)$, then the collection $\{\varphi_k(x)\varphi_\ell(y)\}_{k,\ell\geq 1}$ is an orthonormal basis for $L^2(\mathbb{R}^d \times \mathbb{R}^d)$; the proof of this simple fact is outlined in Exercise 7. As a result

$$K(x,y) \sim \sum_{k,\ell=1}^{\infty} a_{k\ell}\varphi_k(x)\varphi_\ell(y), \qquad \text{with } \textstyle\sum_{k,\ell} |a_{k\ell}|^2 < \infty.$$

We define an operator

$$T_n f(x) = \int_{\mathbb{R}^d} K_n(x,y) f(y) dy, \qquad \text{where } K_n(x,y) = \textstyle\sum_{k,\ell=1}^{n} a_{k\ell}\varphi_k(x)\varphi_\ell(y).$$

Then, each T_n has finite-dimensional range, hence is compact. Moreover,

$$\|K - K_n\|^2_{L^2(\mathbb{R}^d\times\mathbb{R}^d)} = \sum_{k\,\geq\, n \text{ or } \ell\,\geq\, n} |a_{k\ell}|^2 \to 0 \qquad \text{as } n \to \infty.$$

By Proposition 5.5, $\|T - T_n\| \leq \|K - K_n\|_{L^2(\mathbb{R}^d\times\mathbb{R}^d)}$, so we can conclude the proof that T is compact by appealing to Proposition 6.1.

The climax of our efforts regarding compact operators is the infinite-dimensional version of the familiar diagonalization theorem in linear algebra for symmetric matrices. Using a similar terminology, we say that a bounded linear operator T is **symmetric** if $T^* = T$. (These operators are also called "self-adjoint" or "Hermitian.")

Theorem 6.2 (Spectral theorem) *Suppose T is a compact symmetric operator on a Hilbert space $\mathcal{H}$. Then there exists an (orthonormal) basis $\{\varphi_k\}_{k=1}^\infty$ of $\mathcal{H}$ that consists of eigenvectors of T. Moreover, if*

$$T\varphi_k = \lambda_k\varphi_k,$$

then $\lambda_k \in \mathbb{R}$ and $\lambda_k \to 0$ as $k \to \infty$.

Conversely, every operator of the above form is compact and symmetric.

The collection $\{\lambda_k\}$ is called the **spectrum** of T.

Lemma 6.3 *Suppose T is a bounded symmetric linear operator on a Hilbert space $\mathcal{H}$.*

(i) *If λ is an eigenvalue of T, then λ is real.*

(ii) *If f_1 and f_2 are eigenvectors corresponding to two distinct eigenvalues, then f_1 and f_2 are orthogonal.*

Proof. To prove (i), we first choose a non-zero eigenvector f such that $T(f) = \lambda f$. Since T is symmetric (that is, $T = T^*$), we find that

$$\lambda(f, f) = (Tf, f) = (f, Tf) = (f, \lambda f) = \overline{\lambda}(f, f),$$

where we have used in the last equality the fact that the inner product is conjugate linear in the second variable. Since $f \neq 0$, we must have $\lambda = \overline{\lambda}$ and hence $\lambda \in \mathbb{R}$.

For (ii), suppose f_1 and f_2 have eigenvalues λ_1 and λ_2, respectively. By the previous argument both λ_1 and λ_2 are real, and we note that

$$\begin{aligned}
\lambda_1(f_1, f_2) &= (\lambda_1 f_1, f_2) \\
&= (Tf_1, f_2) \\
&= (f_1, Tf_2) \\
&= (f_1, \lambda_2 f_2) \\
&= \lambda_2(f_1, f_2).
\end{aligned}$$

Since by assumption $\lambda_1 \neq \lambda_2$ we must have $(f_1, f_2) = 0$ as desired.

For the next lemma note that every non-zero element of the null-space of $T - \lambda I$ is an eigenvector with eigenvalue λ.

Lemma 6.4 *Suppose T is compact, and $\lambda \neq 0$. Then the dimension of the null space of $T - \lambda I$ is finite. Moreover, the eigenvalues of T form at most a denumerable set $\lambda_1, \ldots, \lambda_k, \ldots$, with $\lambda_k \to 0$ as $k \to \infty$. More specifically, for each $\mu > 0$, the linear space spanned by the eigenvectors corresponding to the eigenvalues λ_k with $|\lambda_k| > \mu$ is finite-dimensional.*

Proof. Let V_λ denote the null-space of $T - \lambda I$, that is, the eigenspace of T corresponding to λ. If V_λ is not finite-dimensional, there exists a countable sequence of orthonormal vectors $\{\varphi_k\}$ in V_λ. Since T is compact, there exists a subsequence $\{\varphi_{n_k}\}$ such that $T(\varphi_{n_k})$ converges.

But since $T(\varphi_{n_k}) = \lambda\varphi_{n_k}$ and $\lambda \neq 0$, we conclude that φ_{n_k} converges, which is a contradiction since $\|\varphi_{n_k} - \varphi_{n_{k'}}\|^2 = 2$ if $k \neq k'$.

The rest of the lemma follows if we can show that for each $\mu > 0$, there are only finitely many eigenvalues whose absolute values are greater than μ. We argue again by contradiction. Suppose there are infinitely many distinct eigenvalues whose absolute values are greater than μ, and let $\{\varphi_k\}$ be a corresponding sequence of eigenvectors. Since the eigenvalues are distinct, we know from the previous lemma that $\{\varphi_k\}$ is orthogonal, and after normalization, we may assume that this set of eigenvectors is orthonormal. One again, since T is compact, we may find a subsequence so that $T(\varphi_{n_k})$ converges, and since

$$T(\varphi_{n_k}) = \lambda_{n_k}\varphi_{n_k}$$

the fact that $|\lambda_{n_k}| > \mu$ leads to a contradiction, since $\{\varphi_k\}$ is an orthonormal set and thus $\|\lambda_{n_k}\varphi_{n_k} - \lambda_{n_j}\varphi_{n_j}\|^2 = \lambda_{n_k}^2 + \lambda_{n_j}^2 \geq 2\mu^2$.

Lemma 6.5 *Suppose $T \neq 0$ is compact and symmetric. Then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. By the observation (7) made earlier, either

$$\|T\| = \sup\{(Tf, f) : \ \|f\| = 1\} \quad \text{or} \quad -\|T\| = \inf\{(Tf, f) : \ \|f\| = 1\}.$$

We assume the first case, that is,

$$\lambda = \|T\| = \sup\{(Tf, f) : \ \|f\| = 1\},$$

and prove that λ is an eigenvalue of T. (The proof of the other case is similar.)

We pick a sequence $\{f_n\} \subset \mathcal{H}$ such that $\|f_n\| = 1$ and $(Tf_n, f_n) \to \lambda$. Since T is compact, we may assume also (by passing to a subsequence of $\{f_n\}$ if necessary) that $\{Tf_n\}$ converges to a limit $g \in \mathcal{H}$. We claim that g is an eigenvector of T with eigenvalue λ. To see this, we first observe that $Tf_n - \lambda f_n \to 0$ because

$$\begin{aligned}\|Tf_n - \lambda f_n\|^2 &= \|Tf_n\|^2 - 2\lambda(Tf_n, f_n) + \lambda^2\|f_n\|^2 \\ &\leq \|T\|^2\|f_n\|^2 - 2\lambda(Tf_n, f_n) + \lambda^2\|f_n\|^2 \\ &\leq 2\lambda^2 - 2\lambda(Tf_n, f_n) \to 0.\end{aligned}$$

Since $Tf_n \to g$, we must have $\lambda f_n \to g$, and since T is continuous, this implies that $\lambda Tf_n \to Tg$. This proves that $\lambda g = Tg$. Finally, we must

have $g \neq 0$, for otherwise $\|T_n f_n\| \to 0$, hence $(Tf_n, f_n) \to 0$, and $\lambda = \|T\| = 0$, which is a contradiction.

We are now equipped with the necessary tools to prove the spectral theorem. Let $\mathcal{S}$ denote the closure of the linear space spanned by all eigenvectors of T. By Lemma 6.5, the space $\mathcal{S}$ is non-empty. The goal is to prove that $\mathcal{S} = \mathcal{H}$. If not, then since

$$\mathcal{S} \oplus \mathcal{S}^{\perp} = \mathcal{H}, \tag{9}$$

$\mathcal{S}^{\perp}$ would be non-empty. We will have reached a contradiction once we show that $\mathcal{S}^{\perp}$ contains an eigenvector of T. First, we note that T respects the decomposition (9). In other words, if $f \in \mathcal{S}$ then $Tf \in \mathcal{S}$, which follows from the definitions. Also, if $g \in \mathcal{S}^{\perp}$ then $Tg \in \mathcal{S}^{\perp}$. This is because T is symmetric and maps $\mathcal{S}$ to itself, and hence

$$(Tg, f) = (g, Tf) = 0 \quad \text{whenever } g \in \mathcal{S}^{\perp} \text{ and } f \in \mathcal{S}.$$

Now consider the operator T_1, which by definition is the restriction of T to the subspace $S^{\perp}$. The closed subspace $S^{\perp}$ inherits its Hilbert space structure from $\mathcal{H}$. We see immediately that T_1 is also a compact and symmetric operator on this Hilbert space. Moreover, if $S^{\perp}$ is non-empty, the lemma implies that T_1 has a non-zero eigenvector in $S^{\perp}$. This eigenvector is clearly also an eigenvector of T, and therefore a contradiction is obtained. This concludes the proof of the spectral theorem.

Some comments about Theorem 6.2 are in order. If in its statement we drop either of the two assumptions (the compactness or symmetry of T), then T may have no eigenvectors. (See Exercises 32 and 33.) However, when T is a general bounded linear transformation which is symmetric, there is an appropriate extension of the spectral theorem that holds for it. Its formulation and proof require further ideas that are deferred to Chapter 6.

7 Exercises

1. Show that properties (i) and (ii) in the definition of a Hilbert space (Section 2) imply property (iii): the Cauchy-Schwarz inequality $|(f, g)| \leq \|f\| \cdot \|g\|$ and the triangle inequality $\|f + g\| \leq \|f\| + \|g\|$.

[Hint: For the first inequality, consider $(f + \lambda g, f + \lambda g)$ as a positive quadratic function of λ. For the second, write $\|f + g\|^2$ as $(f + g, f + g)$.]

2. In the case of equality in the Cauchy-Schwarz inequality we have the following. If $|(f, g)| = \|f\|\,\|g\|$ and $g \neq 0$, then $f = cg$ for some scalar c.

[Hint: Assume $\|f\| = \|g\| = 1$ and $(f, g) = 1$. Then $f - g$ and g are orthogonal, while $f = f - g + g$. Thus $\|f\|^2 = \|f - g\|^2 + \|g\|^2$.]

3. Note that $\|f + g\|^2 = \|f\|^2 + \|g\|^2 + 2\text{Re}(f, g)$ for any pair of elements in a Hilbert space $\mathcal{H}$. As a result, verify the identity $\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2)$.

4. Prove from the definition that $\ell^2(\mathbb{Z})$ is complete and separable.

5. Establish the following relations between $L^2(\mathbb{R}^d)$ and $L^1(\mathbb{R}^d)$:

(a) Neither the inclusion $L^2(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$ nor the inclusion $L^1(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ is valid.

(b) Note, however, that if f is supported on a set E of finite measure and if $f \in L^2(\mathbb{R}^d)$, applying the Cauchy-Schwarz inequality to $f\chi_E$ gives $f \in L^1(\mathbb{R}^d)$, and

$$\|f\|_{L^1(\mathbb{R}^d)} \leq m(E)^{1/2}\|f\|_{L^2(\mathbb{R}^d)}.$$

(c) If f is bounded ($|f(x)| \leq M$), and $f \in L^1(\mathbb{R}^d)$, then $f \in L^2(\mathbb{R}^d)$ with

$$\|f\|_{L^2(\mathbb{R}^d)} \leq M^{1/2}\|f\|^{1/2}_{L^1(\mathbb{R}^d)}.$$

[Hint: For (a) consider $f(x) = |x|^{-\alpha}$, when $|x| \leq 1$ or when $|x| > 1$.]

6. Prove that the following are dense subspaces of $L^2(\mathbb{R}^d)$.

(a) The simple functions.

(b) The continuous functions of compact support.

7. Suppose $\{\varphi_k\}_{k=1}^{\infty}$ is an orthonormal basis for $L^2(\mathbb{R}^d)$. Prove that the collection $\{\varphi_{k,j}\}_{1\leq k,j<\infty}$ with $\varphi_{k,j}(x, y) = \varphi_k(x)\varphi_j(y)$ is an orthonormal basis of $L^2(\mathbb{R}^d \times \mathbb{R}^d)$.
[Hint: First verify that the $\{\varphi_{k,j}\}$ are orthonormal, by Fubini's theorem. Next, for each j consider $F_j(x) = \int_{\mathbb{R}^d} F(x, y)\overline{\varphi_j(y)}\, dy$. If one assumes that $(F, \varphi_{k,j}) = 0$ for all j, then $\int F_j(x)\overline{\varphi_k(x)}\, dx = 0$.]

8. Let $\eta(t)$ be a fixed strictly positive continuous function on $[a, b]$. Define $\mathcal{H}_\eta = L^2([a, b], \eta)$ to be the space of all measurable functions f on $[a, b]$ such that

$$\int_a^b |f(t)|^2\eta(t)\, dt < \infty.$$

Define the inner product on $\mathcal{H}_\eta$ by

$$(f, g)_\eta = \int_a^b f(t)\overline{g(t)}\eta(t)\, dt.$$

(a) Show that $\mathcal{H}_\eta$ is a Hilbert space, and that the mapping $U : f \mapsto \eta^{1/2} f$ gives a unitary correspondence between $\mathcal{H}_\eta$ and the usual space $L^2([a, b])$.

(b) Generalize this to the case when η is not necessarily continuous.

9. Let $\mathcal{H}_1 = L^2([-\pi, \pi])$ be the Hilbert space of functions $F(e^{i\theta})$ on the unit circle with inner product $(F, G) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{i\theta}) \overline{G(e^{i\theta})}\, d\theta$. Let $\mathcal{H}_2$ be the space $L^2(\mathbb{R})$. Using the mapping

$$x \mapsto \frac{i - x}{i + x}$$

of $\mathbb{R}$ to the unit circle, show that:

(a) The correspondence $U : F \to f$, with

$$f(x) = \frac{1}{\pi^{1/2}(i + x)} F\left(\frac{i - x}{i + x}\right)$$

gives a unitary mapping of $\mathcal{H}_1$ to $\mathcal{H}_2$.

(b) As a result,

$$\left\{ \frac{1}{\pi^{1/2}} \left(\frac{i - x}{i + x}\right)^n \frac{1}{i + x} \right\}_{n=-\infty}^{\infty}$$

is an orthonormal basis of $L^2(\mathbb{R})$.

10. Let $\mathcal{S}$ denote a subspace of a Hilbert space $\mathcal{H}$. Prove that $(\mathcal{S}^\perp)^\perp$ is the smallest closed subspace of $\mathcal{H}$ that contains $\mathcal{S}$.

11. Let P be the orthogonal projection associated with a closed subspace $\mathcal{S}$ in a Hilbert space $\mathcal{H}$, that is,

$$P(f) = f \text{ if } f \in \mathcal{S} \quad \text{and} \quad P(f) = 0 \text{ if } f \in \mathcal{S}^\perp.$$

(a) Show that $P^2 = P$ and $P^* = P$.

(b) Conversely, if P is any bounded operator satisfying $P^2 = P$ and $P^* = P$, prove that P is the orthogonal projection for some closed subspace of $\mathcal{H}$.

(c) Using P, prove that if $\mathcal{S}$ is a closed subspace of a separable Hilbert space, then $\mathcal{S}$ is also a separable Hilbert space.

12. Let E be a measurable subset of $\mathbb{R}^d$, and suppose $\mathcal{S}$ is the subspace of $L^2(\mathbb{R}^d)$ of functions that vanish for a.e. $x \notin E$. Show that the orthogonal projection P on $\mathcal{S}$ is given by $P(f) = \chi_E \cdot f$, where χ_E is the characteristic function of E.

13. Suppose P_1 and P_2 are a pair of orthogonal projections on S_1 and S_2, respectively. Then P_1P_2 is an orthogonal projection if and only if P_1 and P_2 commute, that is, $P_1P_2 = P_2P_1$. In this case, P_1P_2 projects onto $S_1 \cap S_2$.

14. Suppose $\mathcal{H}$ and $\mathcal{H}'$ are two completions of a pre-Hilbert space $\mathcal{H}_0$. Show that there is a unitary mapping from $\mathcal{H}$ to $\mathcal{H}'$ that is the identity on $\mathcal{H}_0$.

[Hint: If $f \in \mathcal{H}$, pick a Cauchy sequence $\{f_n\}$ in $\mathcal{H}_0$ that converges to f in $\mathcal{H}$. This sequence will also converge to an element f' in $\mathcal{H}'$. The mapping $f \mapsto f'$ gives the required unitary mapping.]

15. Let T be any linear transformation from $\mathcal{H}_1$ to $\mathcal{H}_2$. If we suppose that $\mathcal{H}_1$ is finite-dimensional, then T is automatically bounded. (If $\mathcal{H}_1$ is not assumed to be finite-dimensional this may fail; see Problem 1 below.)

16. Let $F_0(z) = 1/(1-z)^i$.

(a) Verify that $|F_0(z)| \leq e^{\pi/2}$ in the unit disc, but that $\lim_{r\to 1} F_0(r)$ does not exist.

[Hint: Note that $|F_0(r)| = 1$ and $F_0(r)$ oscillates between ± 1 infinitely often as $r \to 1$.]

(b) Let $\{\alpha_n\}_{n=1}^{\infty}$ be an enumeration of the rationals, and let

$$F(z) = \sum_{j=1}^{\infty} \delta^j F_0(ze^{-i\alpha_j}),$$

where δ is sufficiently small. Show that $\lim_{r\to 1} F(re^{i\theta})$ fails to exist whenever $\theta = \alpha_j$, and hence F fails to have a radial limit for a dense set of points on the unit circle.

17. Fatou's theorem can be generalized by allowing a point to approach the boundary in larger regions, as follows.

For each $0 < s < 1$ and point z on the unit circle, consider the region $\Gamma_s(z)$ defined as the smallest closed convex set that contains z and the closed disc $D_s(0)$. In other words, $\Gamma_s(z)$ consists of all lines joining z with points in $D_s(0)$. Near the point z, the region $\Gamma_s(z)$ looks like a triangle. See Figure 2.

We say that a function F defined in the open unit disc has a **non-tangential limit** at a point z on the circle, if for every $0 < s < 1$, the limit

$$\lim_{\substack{w \to z \\ w \in \Gamma_s(z)}} F(w)$$

exists.

Prove that if F is holomorphic and bounded on the open unit disc, then F has a non-tangential limit for almost every point on the unit circle.

[Hint: Show that the Poisson integral of a function f has non-tangential limits at every point of the Lebesgue set of f.]

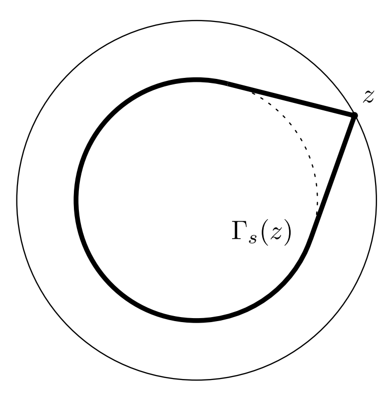


Figure 2. The region $\Gamma_s(z)$

18. Let $\mathcal{H}$ denote a Hilbert space, and $\mathcal{L}(\mathcal{H})$ the vector space of all bounded linear operators on $\mathcal{H}$. Given $T \in \mathcal{L}(\mathcal{H})$, we define the operator norm

$$\|T\| = \inf\{B : \|Tv\| \leq B\|v\|, \quad \text{for all } v \in \mathcal{H}\}.$$

(a) Show that $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$ whenever $T_1, T_2 \in \mathcal{L}(\mathcal{H})$.

(b) Prove that

$$d(T_1, T_2) = \|T_1 - T_2\|$$

defines a metric on $\mathcal{L}(\mathcal{H})$.

(c) Show that $\mathcal{L}(\mathcal{H})$ is complete in the metric d.

19. If T is a bounded linear operator on a Hilbert space, prove that

$$\|TT^*\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2.$$

20. Suppose $\mathcal{H}$ is an infinite-dimensional Hilbert space. We have seen an example of a sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, but for which no subsequence of $\{f_n\}$ converges in $\mathcal{H}$. However, show that for any sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, there exist $f \in \mathcal{H}$ and a subsequence $\{f_{n_k}\}$ such that for all $g \in \mathcal{H}$, one has

$$\lim_{k \to \infty} (f_{n_k}, g) = (f, g).$$

One says that $\{f_{n_k}\}$ **converges weakly** to f.

[Hint: Let g run through a basis for $\mathcal{H}$, and use a diagonalization argument. One can then define f by giving its series expansion with respect to the chosen basis.]

21. There are several senses in which a sequence of bounded operators $\{T_n\}$ can converge to a bounded operator T (in a Hilbert space $\mathcal{H}$). First, there is convergence in the norm, that is, $\|T_n - T\| \to 0$, as $n \to \infty$. Next, there is a weaker convergence, which happens to be called **strong convergence**, that requires that $T_n f \to Tf$, as $n \to \infty$, for every vector $f \in \mathcal{H}$. Finally, there is **weak convergence** (see also Exercise 20) that requires $(T_n f, g) \to (Tf, g)$ for every pair of vectors $f, g \in \mathcal{H}$.

(a) Show by examples that weak convergence does not imply strong convergence, nor does strong convergence imply convergence in the norm.

(b) Show that for any bounded operator T there is a sequence $\{T_n\}$ of bounded operators of finite rank so that $T_n \to T$ strongly as $n \to \infty$.

22. An operator T is an **isometry** if $\|Tf\| = \|f\|$ for all $f \in \mathcal{H}$.

(a) Show that if T is an isometry, then $(Tf, Tg) = (f, g)$ for every $f, g \in \mathcal{H}$. Prove as a result that $T^*T = I$.

(b) If T is an isometry and T is surjective, then T is unitary and $TT^* = I$.

(c) Give an example of an isometry that is not unitary.

(d) Show that if T^*T is unitary then T is an isometry.

[Hint: Use the fact that $(Tf, Tf) = (f, f)$ for f replaced by $f \pm g$ and $f \pm ig$.]

23. Suppose $\{T_k\}$ is a collection of bounded operators on a Hilbert space $\mathcal{H}$, with $\|T_k\| \leq 1$ for all k. Suppose also that

$$T_k T_j^* = T_k^* T_j = 0 \qquad \text{for all } k \neq j.$$

Let $S_N = \sum_{k=-N}^{N} T_k$.

Show that $S_N(f)$ converges as $N \to \infty$, for every $f \in \mathcal{H}$. If $T(f)$ denotes the limit, prove that $\|T\| \leq 1$.

A generalization is given in Problem 8* below.

[Hint: Consider first the case when only finitely many of the T_k are non-zero, and note that the ranges of the T_k are mutually orthogonal.]

24. Let $\{e_k\}_{k=1}^{\infty}$ denote an orthonormal set in a Hilbert space $\mathcal{H}$. If $\{c_k\}_{k=1}^{\infty}$ is a sequence of positive real numbers such that $\sum c_k^2 < \infty$, then the set

$$A = \{\sum_{k=1}^{\infty} a_k e_k : \ |a_k| \leq c_k\}$$

is compact in $\mathcal{H}$.

25. Suppose T is a bounded operator that is diagonal with respect to a basis $\{\varphi_k\}$, with $T\varphi_k = \lambda_k \varphi_k$. Then T is compact if and only if $\lambda_k \to 0$.

[Hint: If $\lambda_k \to 0$, then note that $\|P_n T - T\| \to 0$, where P_n is the orthogonal projection on the subspace spanned by $\varphi_1, \varphi_2, \ldots, \varphi_n$.]

26. Suppose w is a measurable function on $\mathbb{R}^d$ with $0 < w(x) < \infty$ for a.e. x, and K is a measurable function on $\mathbb{R}^{2d}$ that satisfies:

(i) $\displaystyle\int_{\mathbb{R}^d} |K(x,y)| w(y)\, dy \leq A w(x)$ for almost every $x \in \mathbb{R}^d$, and

(ii) $\displaystyle\int_{\mathbb{R}^d} |K(x,y)| w(x)\, dx \leq A w(y)$ for almost every $y \in \mathbb{R}^d$.

Prove that the integral operator defined by

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\, dy, \qquad x \in \mathbb{R}^d$$

is bounded on $L^2(\mathbb{R}^d)$ with $\|T\| \leq A$.

Note as a special case that if $\int |K(x,y)|\, dy \leq A$ for all x, and $\int |K(x,y)|\, dx \leq A$ for all y, then $\|T\| \leq A$.

[Hint: Show that if $f \in L^2(\mathbb{R}^d)$, then

$$\int |K(x,y)|\, |f(y)|\, dy \leq A^{1/2} w(x)^{1/2} \left[\int |K(x,y)|\, |f(y)|^2 w(y)^{-1}\, dy \right]^{1/2}.]$$

27. Prove that the operator

$$Tf(x) = \frac{1}{\pi} \int_0^\infty \frac{f(y)}{x+y}\, dy$$

is bounded on $L^2(0,\infty)$ with norm $\|T\| \leq 1$.

[Hint: Use Exercise 26 with an appropriate w.]

28. Suppose $\mathcal{H} = L^2(B)$, where B is the unit ball in $\mathbb{R}^d$. Let $K(x,y)$ be a measurable function on $B \times B$ that satisfies $|K(x,y)| \leq A|x-y|^{-d+\alpha}$ for some $\alpha > 0$, whenever $x, y \in B$. Define

$$Tf(x) = \int_B K(x,y) f(y) dy.$$

(a) Prove that T is a bounded operator on $\mathcal{H}$.

(b) Prove that T is compact.

(c) Note that T is a Hilbert-Schmidt operator if and only if $\alpha > d/2$.

[Hint: For (b), consider the operators T_n associated with the truncated kernels $K_n(x,y) = K(x,y)$ if $|x-y| \geq 1/n$ and 0 otherwise. Show that each T_n is compact, and that $\|T_n - T\| \to 0$ as $n \to \infty$.]

29. Let T be a compact operator on a Hilbert space $\mathcal{H}$, and assume $\lambda \neq 0$.

(a) Show that the range of $\lambda I - T$ defined by

$$\{g \in \mathcal{H} : g = (\lambda I - T)f, \text{ for some } f \in \mathcal{H}\}$$

is closed. [Hint: Suppose $g_j \to g$, where $g_j = (\lambda I - T)f_j$. Let V_λ denote the eigenspace of T corresponding to λ, that is, the kernel of $\lambda I - T$. Why can one assume that $f_j \in V_\lambda^\perp$? Under this assumption prove that $\{f_j\}$ is a bounded sequence.]

(b) Show by example that this may fail when $\lambda = 0$.

(c) Show that the range of $\lambda I - T$ is all of $\mathcal{H}$ if and only if the null-space of $\overline{\lambda} I - T^*$ is trivial.

30. Let $\mathcal{H} = L^2([-\pi, \pi])$ with $[-\pi, \pi]$ identified as the unit circle. Fix a bounded sequence $\{\lambda_n\}_{n=-\infty}^{\infty}$ of complex numbers, and define an operator Tf by

$$Tf(x) \sim \sum_{n=-\infty}^{\infty} \lambda_n a_n e^{inx} \qquad \text{whenever} \qquad f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$

Such an operator is called a **Fourier multiplier operator**, and the sequence $\{\lambda_n\}$ is called the **multiplier sequence**.

(a) Show that T is a bounded operator on $\mathcal{H}$ and $\|T\| = \sup_n |\lambda_n|$.

(b) Verify that T commutes with translations, that is, if we define $\tau_h(x) = f(x - h)$ then

$$T \circ \tau_h = \tau_h \circ T \qquad \text{for every } h \in \mathbb{R}.$$

(c) Conversely, prove that if T is any bounded operator on $\mathcal{H}$ that commutes with translations, then T is a Fourier multiplier operator. [Hint: Consider $T(e^{inx})$.]

31. Consider a version of the sawtooth function defined on $[-\pi, \pi)$ by[5]

$$K(x) = i(sgn(x)\pi - x),$$

and extended to $\mathbb{R}$ with period 2π. Suppose $f \in L^1([-\pi, \pi])$ is extended to $\mathbb{R}$ with period 2π, and define

$$\begin{aligned} Tf(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} K(x - y) f(y)\, dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} K(y) f(x - y)\, dy. \end{aligned}$$

[5]The symbol $sgn(x)$ denotes the sign function: it equals 1 or -1 if x is positive or negative respectively, and 0 if $x = 0$.

(a) Show that $F(x) = Tf(x)$ is absolutely continuous, and if $\int_{-\pi}^{\pi} f(y)dy = 0$, then $F'(x) = if(x)$ a.e. x.

(b) Show that the mapping $f \mapsto Tf$ is compact and symmetric on $L^2([-\pi, \pi])$.

(c) Prove that $\varphi(x) \in L^2([-\pi, \pi])$ is an eigenfunction for T if and only if $\varphi(x)$ is (up to a constant multiple) equal to e^{inx} for some integer $n \neq 0$ with eigenvalue $1/n$, or $\varphi(x) = 1$ with eigenvalue 0.

(d) Show as a result that $\{e^{inx}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2([-\pi, \pi])$.

Note that in Book I, Chapter 2, Exercise 8, it is shown that the Fourier series of K is

$$K(x) \sim \sum_{n\neq 0} \frac{e^{inx}}{n}.$$

32. Consider the operator $T : L^2([0, 1]) \to L^2([0, 1])$ defined by

$$T(f)(t) = tf(t).$$

(a) Prove that T is a bounded linear operator with $T = T^*$, but that T is not compact.

(b) However, show that T has no eigenvectors.

33. Let $\mathcal{H}$ be a Hilbert space with basis $\{\varphi_k\}_{k=1}^{\infty}$. Verify that the operator T defined by

$$T(\varphi_k) = \frac{1}{k}\,\varphi_{k+1}$$

is compact, but has no eigenvectors.

34. Let K be a Hilbert-Schmidt kernel which is real and symmetric. Then, as we saw, the operator T whose kernel is K is compact and symmetric. Let $\{\varphi_k(x)\}$ be the eigenvectors (with eigenvalues λ_k) that diagonalize T. Then:

(a) $\sum_k |\lambda_k|^2 < \infty$.

(b) $K(x, y) \sim \sum \lambda_k \varphi_k(x)\varphi_k(y)$ is the expansion of K in the basis $\{\varphi_k(x)\varphi_k(y)\}$.

(c) Suppose T is a compact operator which is symmetric. Then T is of Hilbert-Schmidt type if and only if $\sum_n |\lambda_n|^2 < \infty$, where $\{\lambda_n\}$ are the eigenvalues of T counted according to their multiplicities.

35. Let $\mathcal{H}$ be a Hilbert space. Prove the following variants of the spectral theorem.

(a) If T_1 and T_2 are two linear symmetric and compact operators on $\mathcal{H}$ that commute (that is, $T_1T_2 = T_2T_1$), show that they can be diagonalized simultaneously. In other words, there exists an orthonormal basis for $\mathcal{H}$ which consists of eigenvectors for both T_1 and T_2.

(b) A linear operator on $\mathcal{H}$ is **normal** if $TT^* = T^*T$. Prove that if T is normal and compact, then T can be diagonalized.

[Hint: Write $T = T_1 + iT_2$ where T_1 and T_2 are symmetric, compact and commute.]

(c) If U is unitary, and $U = \lambda I - T$, where T is compact, then U can be diagonalized.

8 Problems

1. Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. There exists a linear functional ℓ defined on $\mathcal{H}$ that is not bounded (and hence not continuous).

[Hint: Using the axiom of choice (or one of its equivalent forms), construct an **algebraic basis** of $\mathcal{H}$, $\{e_\alpha\}$; it has the property that every element of $\mathcal{H}$ is uniquely a finite linear combination of the $\{e_\alpha\}$. Select a denumerable collection $\{e_n\}_{n=1}^\infty$, and define ℓ to satisfy the requirement that $\ell(e_n) = n\|e_n\|$ for all $n \in \mathbb{N}$.]

2.* The following is an example of a non-separable Hilbert space. We consider the collection of exponentials $\{e^{i\lambda x}\}$ on $\mathbb{R}$, where λ ranges over the real numbers. Let $\mathcal{H}_0$ denote the space of finite linear combinations of these exponentials. For $f, g \in \mathcal{H}_0$, we define the inner product as

$$(f, g) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x)\overline{g(x)}\, dx.$$

(a) Show that this limit exists, and

$$(f, g) = \sum_{k=1}^{N} a_{\lambda_k}\overline{b_{\lambda_k}}$$

if $f(x) = \sum_{k=1}^N a_{\lambda_k} e^{i\lambda_k x}$ and $g(x) = \sum_{k=1}^N b_{\lambda_k} e^{i\lambda_k x}$.

(b) With this inner product $\mathcal{H}_0$ is a pre-Hilbert space. Notice that $\|f\| \leq \sup_x |f(x)|$, if $f \in \mathcal{H}_0$, where $\|f\|$ denotes the norm $\langle f, f\rangle^{1/2}$. Let $\mathcal{H}$ be the completion of $\mathcal{H}_0$. Then $\mathcal{H}$ is not separable because $e^{i\lambda x}$ and $e^{i\lambda' x}$ are orthonormal if $\lambda \neq \lambda'$.

A continuous function F defined on $\mathbb{R}$ is called **almost periodic** if it is the uniform limit (on $\mathbb{R}$) of elements in $\mathcal{H}_0$. Such functions can be identified with (certain) elements in the completion $\mathcal{H}$: We have $\mathcal{H}_0 \subset AP \subset \mathcal{H}$, where AP denotes the almost periodic functions.

(c) A continuous function F is in AP if for every $\epsilon > 0$ we can find a length $L = L_\epsilon$ such that any interval $I \subset \mathbb{R}$ of length L contains an "almost period" τ satisfying

$$\sup_x |F(x+\tau) - F(x)| < \epsilon.$$

(d) An equivalent characterization is that F is in AP if and only if every sequence $F(x+h_n)$ of translates of F contains a subsequence that converges uniformly.

3. The following is a direct generalization of Fatou's theorem: if $u(re^{i\theta})$ is harmonic in the unit disc and bounded there, then $\lim_{r\to 1} u(re^{i\theta})$ exists for a.e. θ.

[Hint: Let $a_n(r) = \frac{1}{2\pi}\int_0^{2\pi} u(re^{i\theta})e^{-in\theta}\,d\theta$. Then $a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{r^2}a_n(r) = 0$, hence $a_n(r) = A_n r^n + B_n r^{-n}$, $n \neq 0$, and as a result[6] $u(re^{i\theta}) = \sum_{-\infty}^{\infty} a_n r^{|n|} e^{in\theta}$. From this one can proceed as in the proof of Theorem 3.3.]

4.* This problem provides some examples of functions that fail to have radial limits almost everywhere.

(a) At almost every point of the boundary unit circle, the function $\sum_{n=0}^{\infty} z^{2^n}$ fails to have a radial limit.

(b) More generally, suppose $F(z) = \sum_{n=0}^{\infty} a_n z^{2^n}$. Then, if $\sum |a_n|^2 = \infty$ the function F fails to have radial limits at almost every boundary point. However, if $\sum |a_n|^2 < \infty$, then $F \in H^2(\mathbb{D})$, and we know by the proof of Theorem 3.3 that F does have radial limits almost everywhere.

5.* Suppose F is holomorphic in the unit disc, and

$$\sup_{0\le r<1} \frac{1}{2\pi}\int_{-\pi}^{\pi} \log^+ |F(re^{i\theta})|\,d\theta < \infty,$$

where $\log^+ u = \log u$ if $u \geq 1$, and $\log^+ u = 0$ if $u < 1$.

Then $\lim_{r\to 1} F(re^{i\theta})$ exists for almost every θ.

The above condition is satisfied whenever (say)

$$\sup_{0\le r<1} \frac{1}{2\pi}\int_{-\pi}^{\pi} |F(re^{i\theta})|^p\,d\theta < \infty, \qquad \text{for some } p > 0,$$

(since $e^{pu} \geq pu$, $u \geq 0$).

Functions that satisfy the latter condition are said to belong to the Hardy space $H^p(\mathbb{D})$.

6.* If T is compact, and $\lambda \neq 0$, show that

[6] See also Section 5, Chapter 2 in Book I.

(a) $\lambda I - T$ is injective if and only if $\overline{\lambda} I - T^*$ is injective.

(b) $\lambda I - T$ is injective if and only if $\lambda I - T$ is surjective.

This result, known as the **Fredholm alternative**, is often combined with that in Exercise 29.

7. Show that the identity operator on $L^2(\mathbb{R}^d)$ cannot be given as an (absolutely) convergent integral operator. More precisely, if $K(x, y)$ is a measurable function on $\mathbb{R}^d \times \mathbb{R}^d$ with the property that for each $f \in L^2(\mathbb{R}^d)$, the integral $T(f)(x) = \int_{\mathbb{R}^d} K(x, y)f(y)\, dy$ converges for almost every x, then $T(f) \neq f$ for some f.

[Hint: Prove that otherwise for any pair of disjoint balls B_1 and B_2 in $\mathbb{R}^d$, we would have that $K(x, y) = 0$ for a.e. $(x, y) \in B_1 \times B_2$.]

8.* Suppose $\{T_k\}$ is a collection of bounded opeartors on a Hilbert space $\mathcal{H}$. Assume that

$$\|T_k T_j^*\| \leq a_{k-j} \qquad \text{and} \qquad \|T_k^* T_j\| \leq a_{k-j}^*,$$

for positive constants $\{a_n\}$ with the property that $\sum_{-\infty}^{\infty} a_n = A < \infty$. Then $S_N(f)$ converges as $N \to \infty$, for every $f \in \mathcal{H}$, with $S_N = \sum_{-N}^{N} T_k$. Moreover, $T = \lim_{N\to\infty} S_N$ satisfies $\|T\| \leq A$.

9. A discussion of a class of regular Sturm-Liouville operators follows. Other special examples are given in the problems below.

Suppose $[a, b]$ is a bounded interval, and L is defined on functions f that are twice continuously differentiable in $[a, b]$ (we write, $f \in C^2([a, b])$) by

$$L(f)(x) = \frac{d^2 f}{dx^2} - q(x)f(x).$$

Here the function q is continuous and real-valued on $[a, b]$, and we assume for simplicity that q is non-negative. We say that $\varphi \in C^2([a, b])$ is an **eigenfunction** of L with eigenvalue μ if $L(\varphi) = \mu\varphi$, under the assumption that φ satisfies the boundary conditions $\varphi(a) = \varphi(b) = 0$. Then one can show:

(a) The eigenvalues μ are strictly negative, and the eigenspace corresponding to each eigenvalue is one-dimensional.

(b) Eigenvectors corresponding to distinct eigenvalues are orthogonal in $L^2([a, b])$.

(c) Let $K(x, y)$ be the "Green's kernel" defined as follows. Choose $\varphi_-(x)$ to be a solution of $L(\varphi_-) = 0$, with $\varphi_-(a) = 0$ but $\varphi_-'(a) \neq 0$. Similarly, choose $\varphi_+(x)$ to be a solution of $L(\varphi_+) = 0$ with $\varphi_+(b) = 0$, but $\varphi_+'(b) \neq 0$. Let $w = \varphi_+'(x)\varphi_-(x) - \varphi_-'(x)\varphi_+(x)$, be the "Wronskian" of these solutions, and note that w is a non-zero constant.

Set

$$K(x, y) = \begin{cases} \frac{\varphi_-(x)\varphi_+(y)}{w} & \text{if } a \leq x \leq y \leq b, \\ \frac{\varphi_+(x)\varphi_-(y)}{w} & \text{if } a \leq y \leq x \leq b. \end{cases}$$

Then the operator T defined by

$$T(f)(x) = \int_a^b K(x,y)f(y)\,dy$$

is a Hilbert-Schmidt operator, and hence compact. It is also symmetric. Moreover, whenever f is continuous on $[a,b]$, Tf is of class $C^2([a,b])$ and

$$L(Tf) = f.$$

(d) As a result, each eigenvector of T (with eigenvalue λ) is an eigenvector of L (with eigenvalue $\mu = 1/\lambda$). Hence Theorem 6.2 proves the completeness of the orthonormal set arising from normalizing the eigenvectors of L.

10.* Let L be defined on $C^2([-1,1])$ by

$$L(f)(x) = (1-x^2)\frac{d^2 f}{dx^2} - 2x\frac{df}{dx}.$$

If φ_n is the nth Legendre polynomial, given by

$$\varphi_n(x) = \left(\frac{d}{dx}\right)^n (1-x^2)^n, \qquad n = 0,1,2,\ldots,$$

then $L\varphi_n = -n(n+1)\varphi_n$.

When normalized the φ_n form an orthonormal basis of $L^2([-1,1])$ (see also Problem 2, Chapter 3 in Book I, where φ_n is denoted by L_n.)

11.* The **Hermite functions** $h_k(x)$ are defined by the generating identity

$$\sum_{k=0}^{\infty} h_k(x)\frac{t^k}{k!} = e^{-(x^2/2-2tx+t^2)}.$$

(a) They satisfy the "creation" and "annihilation" identities $\left(x - \frac{d}{dx}\right) h_k(x) = h_{k+1}(x)$ and $\left(x + \frac{d}{dx}\right) h_k(x) = h_{k-1}(x)$ for $k \geq 0$ where $h_{-1}(x) = 0$. Note that $h_0(x) = e^{-x^2/2}$, $h_1(x) = 2xe^{-x^2/2}$, and more generally $h_k(x) = P_k(x)e^{-x^2/2}$, where P_k is a polynomial of degree k.

(b) Using (a) one sees that the h_k are eigenvectors of the operator $L = -d^2/dx^2 + x^2$, with $L(h_k) = \lambda_k h_k$, where $\lambda_k = 2k+1$. One observes that these functions are mutually orthogonal. Since

$$\int_{\mathbb{R}} [h_k(x)]^2\,dx = \pi^{1/2}2^k k! = c_k,$$

we can normalize them obtaining a orthonormal sequence $\{H_k\}$, with $H_k = c_k^{-1/2}h_k$. This sequence is complete in $L^2(\mathbb{R}^d)$ since $\int_{\mathbb{R}} fH_k\,dx = 0$ for all k implies $\int_{-\infty}^{\infty} f(x)e^{-\frac{x^2}{2}+2tx}\,dx = 0$ for all $t \in \mathbb{C}$.

(c) Suppose that $K(x,y) = \sum_{k=0}^{\infty} \frac{H_k(x)H_k(y)}{\lambda_k}$, and also $F(x) = T(f)(x) = \int_{\mathbb{R}} K(x,y)f(y)\,dy$. Then T is a symmetric Hilbert-Schmidt operator, and if $f \sim \sum_{k=0}^{\infty} a_k H_k$, then $F \sim \sum_{k=0}^{\infty} \frac{a_k}{\lambda_k} H_k$.

One can show on the basis of (a) and (b) that whenever $f \in L^2(\mathbb{R})$, not only is $F \in L^2(\mathbb{R})$, but also $x^2 F(x) \in L^2(\mathbb{R})$. Moreover, F can be corrected on a set of measure zero, so it is continuously differentiable, F' is absolutely continuous, and $F'' \in L^2(\mathbb{R})$. Finally, the operator T is the inverse of L in the sense that

$$LT(f) = LF = -F'' + x^2 F = f \qquad \text{for every } f \in L^2(\mathbb{R}).$$

(See also Problem 7* in Chapter 5 of Book I.)

5 Hilbert Spaces: Several Examples

> What is the difference between a mathematician and a physicist? It is this: To a mathematician all Hilbert spaces are the same; for a physicist, however, it is their different realizations that really matter.
>
> *Attributed to E. Wigner,* ca. 1960

Hilbert spaces arise in a large number of different contexts in analysis. Although it is a truism that all (infinite-dimensional) Hilbert spaces are the same, it is in fact their varied and distinct realizations and separate applications that make them of such interest in mathematics. We shall illustrate this via several examples.

To begin with, we consider the Plancherel formula and the resulting unitary character of the Fourier transform. The relevance of these ideas to complex analysis is then highlighted by the study of holomorphic functions in a half-space that belong to the Hardy space H^2. That function space itself is another interesting realization of a Hilbert space. The considerations here are analogous to the ideas that led us to Fatou's theorem for the unit disc, but are of a more involved character.

We next see how complex analysis and the Fourier transform combine to guarantee the existence of solutions to linear partial differential equations with constant coefficients. The proof relies on a basic L^2 estimate, which once established can be exploited by simple Hilbert space techniques.

Our final example is Dirichlet's principle and its applications to the boundary value problem for harmonic functions. Here the Hilbert space that arises is given by Dirichlet's integral, and the solution is expressed by aid of an appropriate orthogonal projection operator.

1 The Fourier transform on L^2

The Fourier transform of a function f on $\mathbb{R}^d$ is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)\, e^{-2\pi i x\cdot\xi}\, dx, \tag{1}$$

and its attached inversion is given by

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi)\, e^{2\pi i x\cdot\xi}\, d\xi. \tag{2}$$

These formulas have already appeared in several different contexts. We considered first (in Book I) the properties of the Fourier transform in the elementary setting by restricting to functions in the Schwartz class $\mathcal{S}(\mathbb{R}^d)$. The class $\mathcal{S}$ consists of functions f that are smooth (indefinitely differentiable) and such that for each multi-index α and β, the function $x^\alpha(\frac{\partial}{\partial x})^\beta f$ is bounded on $\mathbb{R}^d$.[1] We saw that on this class the Fourier transform is a bijection, that the inversion formula (2) holds, and moreover we have the Plancherel identity

$$\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\, d\xi = \int_{\mathbb{R}^d} |f(x)|^2\, dx. \tag{3}$$

Turning now to more general (in particular, non-continuous) functions, we note that the largest class for which the integral defining $\hat{f}(\xi)$ converges (absolutely) is the space $L^1(\mathbb{R}^d)$. For it, we saw in Chapter 2 that a (relative) inversion formula is valid.

Beyond these particular facts, what we would like here is to reestablish in the general context the symmetry between f and $\hat{f}$ that holds for $\mathcal{S}$. This is where the special role of the Hilbert space $L^2(\mathbb{R}^d)$ enters.

We shall define the Fourier transform on $L^2(\mathbb{R}^d)$ as an extension of its definition on $\mathcal{S}$. For this purpose, we temporarily adopt the notational device of denoting by $\mathcal{F}_0$ and $\mathcal{F}$ the Fourier transform on $\mathcal{S}$ and its extension to L^2, respectively.

The main results we prove are the following.

Theorem 1.1 *The Fourier transform $\mathcal{F}_0$, initially defined on $\mathcal{S}(\mathbb{R}^d)$, has a (unique) extension $\mathcal{F}$ to a unitary mapping of $L^2(\mathbb{R}^d)$ to itself. In particular,*

$$\|\mathcal{F}(f)\|_{L^2(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)}$$

for all $f \in L^2(\mathbb{R}^d)$.

The extension $\mathcal{F}$ will be given by a limiting process: if $\{f_n\}$ is a sequence in the Schwartz space that converges to f in $L^2(\mathbb{R}^d)$, then $\{\mathcal{F}_0(f_n)\}$ will

[1] Recall that $x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_d^{\alpha_d}$ and $(\frac{\partial}{\partial x})^\beta = (\frac{\partial}{\partial x_1})^{\beta_1}\cdots(\frac{\partial}{\partial x_d})^{\beta_d}$, where $\alpha = (\alpha_1,\ldots,\alpha_d)$ and $\beta = (\beta_1,\ldots,\beta_d)$, with α_j and β_j positive integers. The order of α is denoted by $|\alpha|$ and defined to be $\alpha_1 + \cdots + \alpha_d$.

converge to an element in $L^2(\mathbb{R}^d)$ which we will *define* as the Fourier transform of f. To implement this approach we have to see that every L^2 function can be approximated by elements in the Schwartz space.

Lemma 1.2 *The space $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$. In other words, given any $f \in L^2(\mathbb{R}^d)$, there exists a sequence $\{f_n\} \subset \mathcal{S}(\mathbb{R}^d)$ such that*

$$\|f - f_n\|_{L^2(\mathbb{R}^d)} \to 0 \qquad \text{as } n \to \infty.$$

For the proof of the lemma, we fix $f \in L^2(\mathbb{R}^d)$ and $\epsilon > 0$. Then, for each $M > 0$, we define

$$g_M(x) = \begin{cases} f(x) & \text{if } |x| \le M \text{ and } |f(x)| \le M, \\ 0 & \text{otherwise.} \end{cases}$$

Then, $|f(x) - g_M(x)| \le 2|f(x)|$, hence $|f(x) - g_M(x)|^2 \le 4|f(x)|^2$, and since $g_M(x) \to f(x)$ as $M \to \infty$ for almost every x, the dominated convergence theorem guarantees that for some M, we have

$$\|f - g_M\|_{L^2(\mathbb{R}^d)} < \epsilon.$$

We write $g = g_M$, note that this function is bounded and supported on a bounded set, and observe that it now suffices to approximate g by functions in the Schwartz space. To achieve this goal, we use a method called **regularization**, which consists of "smoothing" g by convolving it with an approximation of the identity. Consider a function $\varphi(x)$ on $\mathbb{R}^d$ with the following properties:

(a) φ is smooth (indefinitely differentiable).

(b) φ is supported in the unit ball.

(c) $\varphi \ge 0$.

(d) $\displaystyle\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$.

For instance, one can take

$$\varphi(x) = \begin{cases} c\,e^{-\frac{1}{1-|x|^2}} & \text{if } |x| < 1, \\ 0 & \text{if } |x| \ge 1, \end{cases}$$

where the constant c is chosen so that (d) holds.

Next, we consider the approximation to the identity defined by

$$K_\delta(x) = \delta^{-d}\varphi(x/\delta).$$

The key observation is that $g * K_\delta$ belongs to $\mathcal{S}(\mathbb{R}^d)$, with this convolution in fact bounded and supported on a fixed bounded set, uniformly in δ (assuming for example that $\delta \leq 1$). Indeed, we may write

$$(g * K_\delta)(x) = \int g(y) K_\delta(x-y)\, dy = \int g(x-y) K_\delta(y)\, dy,$$

in view of the identity (6) in Chapter 2. We note that since g is supported on some bounded set and K_δ vanishes outside the ball of radius δ, the function $g * K_\delta$ is supported in some fixed bounded set independent of δ. Also, the function g is bounded by construction, hence

$$\begin{aligned} |(g * K_\delta)(x)| &\leq \int |g(x-y)| K_\delta(y)\, dy \\ &\leq \sup_{z \in \mathbb{R}^d} |g(z)| \int K_\delta(y)\, dy = \sup_{z \in \mathbb{R}^d} |g(z)|, \end{aligned}$$

which shows that $g * K_\delta$ is also uniformly bounded in δ. Moreover, from the first integral expression for $g * K_\delta$ above, one may differentiate under the integral sign to see that $g * K_\delta$ is smooth and all of its derivatives have support in some fixed bounded set.

The proof of the lemma will be complete if we can show that $g * K_\delta$ converges to g in $L^2(\mathbb{R}^d)$. Now Theorem 2.1 in Chapter 3 guarantees that for almost every x, the quantity $|(g * K_\delta)(x) - g(x)|^2$ converges to 0 as δ tends to 0. An application of the bounded convergence theorem (Theorem 1.4 in Chapter 2) yields

$$\|(g * K_\delta) - g\|^2_{L^2(\mathbb{R}^d)} \to 0 \quad \text{as } \delta \to 0.$$

In particular, $\|(g * K_\delta) - g\|_{L^2(\mathbb{R}^d)} < \epsilon$ for an appropriate δ and hence $\|f - g * K_\delta\|_{L^2(\mathbb{R}^d)} < 2\epsilon$, and choosing a sequence of ϵ tending to zero gives the construction of the desired sequence $\{f_n\}$.

For later purposes it is useful to observe that the proof of the above lemma establishes the following assertion: if f belongs to both $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$, then there is a sequence $\{f_n\}$, $f_n \in \mathcal{S}(\mathbb{R}^d)$, that converges to f in both the L^1-norm and the L^2-norm.

Our definition of the Fourier transform on $L^2(\mathbb{R}^d)$ combines the above density of $\mathcal{S}$ with a general "extension principle."

Lemma 1.3 *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ denote Hilbert spaces with norms $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively. Suppose $\mathcal{S}$ is a dense subspace of $\mathcal{H}_1$ and $T_0 : \mathcal{S} \to \mathcal{H}_2$ a linear transformation that satisfies $\|T_0(f)\|_2 \leq c\|f\|_1$ whenever $f \in \mathcal{S}$.*

Then T_0 extends to a (unique) linear transformation $T : \mathcal{H}_1 \to \mathcal{H}_2$ that satisfies $\|T(f)\|_2 \leq c\|f\|_1$ for all $f \in \mathcal{H}_1$.

Proof. Given $f \in \mathcal{H}_1$, let $\{f_n\}$ be a sequence in $\mathcal{S}$ that converges to f, and define

$$T(f) = \lim_{n\to\infty} T_0(f_n),$$

where the limit is taken in $\mathcal{H}_2$. To see that T is well-defined we must verify that the limit exists, and that it is independent of the sequence $\{f_n\}$ used to approximate f. Indeed, for the first point, we note that $\{T(f_n)\}$ is a Cauchy sequence in $\mathcal{H}_2$ because by construction $\{f_n\}$ is Cauchy in $\mathcal{H}_1$, and the inequality verified by T_0 yields

$$\|T_0(f_n) - T_0(f_m)\|_2 \leq c\|f_n - f_m\|_1 \to 0 \quad \text{as } n, m \to \infty;$$

thus $\{T_0(f_n)\}$ is Cauchy, hence converges in $\mathcal{H}_2$.

Second, to justify that the limit is independent of the approximating sequence, let $\{g_n\}$ be another sequence in $\mathcal{S}$ that converges to f in $\mathcal{H}_1$. Then

$$\|T_0(f_n) - T_0(g_n)\|_2 \leq c\|f_n - g_n\|_1,$$

and since $\|f_n - g_n\|_1 \leq \|f_n - f\|_1 + \|f - g_n\|_1$, we conclude that $\{T_0(g_n)\}$ converges to a limit in $\mathcal{H}_2$ that equals the limit of $\{T_0(f_n)\}$.

Finally, we recall that if $f_n \to f$ and $T_0(f_n) \to T(f)$, then $\|f_n\|_1 \to \|f\|_1$ and $\|T_0(f_n)\|_2 \to \|T(f)\|_2$, so in the limit as $n \to \infty$, the inequality $\|T(f)\|_2 \leq c\|f\|_1$ holds for all $f \in \mathcal{H}_1$.

In the present case of the Fourier transform, we apply this lemma with $\mathcal{H}_1 = \mathcal{H}_2 = L^2(\mathbb{R}^d)$ (equipped with the L^2-norm), $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$, and $T_0 = \mathcal{F}_0$ the Fourier transform defined on the Schwartz space. The Fourier transform on $L^2(\mathbb{R}^d)$ is by definition the unique (bounded) extension of $\mathcal{F}_0$ to L^2 guaranteed by Lemma 1.3. Thus if $f \in L^2(\mathbb{R}^d)$ and $\{f_n\}$ is any sequence in $\mathcal{S}(\mathbb{R}^d)$ that converges to f (that is, $\|f - f_n\|_{L^2(\mathbb{R}^d)} \to 0$ as $n \to \infty$), we define the **Fourier transform** of f by

$$\mathcal{F}(f) = \lim_{n\to\infty} \mathcal{F}_0(f_n), \tag{4}$$

where the limit is taken in the L^2 sense. Clearly, the argument in the proof of the lemma shows that in our special case the extension $\mathcal{F}$ continues to satisfy the identity (3):

$$\|\mathcal{F}(f)\|_{L^2(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)} \quad \text{whenever } f \in L^2(\mathbb{R}^d).$$

The fact that $\mathcal{F}$ is invertible on L^2 (and thus $\mathcal{F}$ is a unitary mapping) is also a consequence of the analogous property on $\mathcal{S}(\mathbb{R}^d)$. Recall that on the Schwartz space, $\mathcal{F}_0^{-1}$ is given by formula (2), that is,

$$\mathcal{F}_0^{-1}(g)(x) = \int_{\mathbb{R}^d} g(\xi)e^{2\pi i x\cdot\xi}\, d\xi,$$

and satisfies again the identity $\|\mathcal{F}_0^{-1}(g)\|_{L^2} = \|g\|_{L^2}$. Therefore, arguing in the same fashion as above, we can extend $\mathcal{F}_0^{-1}$ to $L^2(\mathbb{R}^d)$ by a limiting argument. Then, given $f \in L^2(\mathbb{R}^d)$, we choose a sequence $\{f_n\}$ in the Schwartz space so that $\|f - f_n\|_{L^2} \to 0$. We have

$$f_n = \mathcal{F}_0^{-1}\mathcal{F}_0(f_n) = \mathcal{F}_0\mathcal{F}_0^{-1}(f_n),$$

and taking the limit as n tends to infinity, we see that

$$f = \mathcal{F}^{-1}\mathcal{F}(f) = \mathcal{F}\mathcal{F}^{-1}(f),$$

and hence $\mathcal{F}$ is invertible. This concludes the proof of Theorem 1.1.

Some remarks are in order.

(i) Suppose f belongs to both $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$. Are the two definitions of the Fourier transform the same? That is, do we have $\mathcal{F}(f) = \hat{f}$, with $\mathcal{F}(f)$ defined by the limiting process in Theorem 1.1 and $\hat{f}$ defined by the convergent integral (1)? To prove that this is indeed the case we recall that we can approximate f by a sequence $\{f_n\}$ in $\mathcal{S}$ so that $f_n \to f$ both in the L^1-norm and the L^2-norm. Since $\mathcal{F}_0(f_n) = \hat{f}_n$, a passage to the limit gives the desired conclusion. In fact, $\mathcal{F}_0(f_n)$ converges to $\mathcal{F}(f)$ in the L^2-norm, so a subsequence converges to $\mathcal{F}(f)$ almost everywhere; see the analogous statement for L^1 in Corollary 2.3, Chapter 2. Moreover,

$$\sup_{\xi\in\mathbb{R}^d} |\hat{f}_n(\xi) - \hat{f}(\xi)| \leq \|f_n - f\|_{L^1(\mathbb{R}^d)},$$

hence $\hat{f}_n$ converges to $\hat{f}$ everywhere, and the assertion is established.

(ii) The theorem gives a rather abstract definition of the Fourier transform on L^2. In view of what we have just said, we can also define the Fourier transform more concretely as follows. If $f \in L^2(\mathbb{R}^d)$, then

$$\hat{f}(\xi) = \lim_{R\to\infty} \int_{|x|\leq R} f(x)e^{-2\pi i x\cdot\xi}\, dx,$$

where the limit is taken in the L^2-norm. Note in fact that if χ_R denotes the characteristic function of the ball $\{x \in \mathbb{R}^d : \ |x| \leq R\}$, then for each R the function $f\chi_R$ is in both L^1 and L^2, and $f\chi_R \to f$ in the L^2-norm.

(iii) The identity of the various definitions of the Fourier transform discussed above allows us to choose $\hat{f}$ as the preferred notation for the Fourier transform. We adopt this practice in what follows.

2 The Hardy space of the upper half-plane

We will apply the L^2 theory of the Fourier transform to holomorphic functions in the upper half-plane. This leads us to consider the relevant analogues of the Hardy space and Fatou's theorem discussed in the previous chapter.[2] It incidentally provides an answer to the following natural question: What are the functions $f \in L^2(\mathbb{R})$ whose Fourier transforms are supported on the half-line $(0, \infty)$?

Let $\mathbb{R}^2_+ = \{z = x + iy,\ x \in \mathbb{R},\ y > 0\}$ be the upper half-plane. We define the **Hardy space** $H^2(\mathbb{R}^2_+)$ to consist of all functions F analytic in $\mathbb{R}^2_+$ with the property that

$$\sup_{y>0} \int_{\mathbb{R}} |F(x+iy)|^2 \, dx < \infty. \tag{5}$$

We define the corresponding norm, $\|F\|_{H^2(\mathbb{R}^2_+)}$, to be the square root of the quantity (5).

Let us first describe a (typical) example of a function F in $H^2(\mathbb{R}^2_+)$. We start with a function $\hat{F}_0$ that belongs to $L^2(0, \infty)$, and write

$$F(x+iy) = \int_0^\infty \hat{F}_0(\xi) e^{2\pi i \xi z} \, d\xi, \qquad z = x+iy,\ y > 0. \tag{6}$$

(The choice of the particular notation $\hat{F}_0$ will become clearer below.) We claim that for any $\delta > 0$ the integral (6) converges absolutely and uniformly as long as $y \geq \delta$. Indeed, $|\hat{F}_0(\xi) e^{2\pi i \xi z}| = |\hat{F}_0(\xi)| e^{-2\pi \xi y}$, hence by the Cauchy-Schwarz inequality

$$\int_0^\infty |\hat{F}_0(\xi) e^{2\pi i \xi z}| \, d\xi \leq \left(\int_0^\infty |\hat{F}_0(\xi)|^2 d\xi \right)^{1/2} \left(\int_0^\infty e^{-4\pi \xi \delta} d\xi \right)^{1/2},$$

from which the asserted convergence is established. From the uniform convergence it follows that $F(z)$ is holomorphic in the upper half-plane. Moreover, by Plancherel's theorem

$$\int_{\mathbb{R}} |F(x+iy)|^2 \, dx = \int_0^\infty |\hat{F}_0(\xi)|^2 \, e^{-4\pi \xi y} \, d\xi \leq \|\hat{F}_0\|^2_{L^2(0,\infty)},$$

[2] Further motivation and some elementary background material may be found in Theorem 3.5 in Chapter 4 of Book II.

and in fact, by the monotone convergence theorem,

$$\sup_{y>0} \int_{\mathbb{R}} |F(x+iy)|^2 \, dx = \|\hat{F}_0\|^2_{L^2(0,\infty)}.$$

In particular, F belongs to $H^2(\mathbb{R}^2_+)$. The main result we prove next is the converse, that is, every element of the space $H^2(\mathbb{R}^2_+)$ is in fact of the form (6).

Theorem 2.1 *The elements F in $H^2(\mathbb{R}^2_+)$ are exactly the functions given by* (6), *with $\hat{F}_0 \in L^2(0,\infty)$. Moreover*

$$\|F\|_{H^2(\mathbb{R}^2_+)} = \|\hat{F}_0\|_{L^2(0,\infty)}.$$

This shows incidentally that $H^2(\mathbb{R}^2_+)$ is a Hilbert space that is isomorphic to $L^2(0,\infty)$ via the correspondence (6).

The crucial point in the proof of the theorem is the following fact. For any fixed strictly positive y, we let $\hat{F}_y(\xi)$ denote the Fourier transform of the L^2 function $F(x+iy)$, $x \in \mathbb{R}$. Then for any pair of choices of y, y_1 and y_2, we have that

$$\hat{F}_{y_1}(\xi)e^{2\pi y_1 \xi} = \hat{F}_{y_2}(\xi)e^{2\pi y_2 \xi} \quad \text{for a.e. } \xi. \tag{7}$$

To establish this assertion we rely on a useful technical observation.

Lemma 2.2 *If F belongs to $H^2(\mathbb{R}^2_+)$, then F is bounded in any proper half-plane $\{z = x+iy,\ y \geq \delta\}$, where $\delta > 0$.*

To prove this we exploit the mean-value property of holomorphic functions. This property may be stated in two alternative ways. First, in terms of averages over circles,

$$F(\zeta) = \frac{1}{2\pi}\int_0^{2\pi} F(\zeta + re^{i\theta})\, d\theta \quad \text{if } 0 < r \leq \delta. \tag{8}$$

(Note that if ζ lies in the upper half-plane, $\text{Im}(\zeta) > \delta$, then the disc centered at ζ of radius r belongs to $\mathbb{R}^2_+$.) Alternatively, integrating over r, we have the mean-value property in terms of discs,

$$F(\zeta) = \frac{1}{\pi\delta^2}\int_{|z|<\delta} F(\zeta + z)\, dx\, dy, \qquad z = x+iy. \tag{9}$$

These assertions actually hold for harmonic functions in $\mathbb{R}^2$ (see Corollary 7.2, Chapter 3 in Book II for the result about holomorphic functions,

and Lemma 2.8, Chapter 5 in Book I for the case of harmonic functions); later in this chapter we in fact prove the extension of (9) to $\mathbb{R}^d$.

From (9) we see from the Cauchy-Schwarz inequality that

$$|F(\zeta)|^2 \leq \frac{1}{\pi\delta^2}\int_{|z|<\delta} |F(\zeta + z)|^2 \, dx \, dy.$$

Writing $z = x + iy$ and $\zeta = \xi + i\eta$, with $\eta > \delta$, we see that the disc $B_\delta(\zeta)$ of center ζ and radius δ is contained in the strip $\{z + \zeta : \ z = x + iy, \ -\delta < y < \delta\}$, and moreover this strip lies in the half-plane $\mathbb{R}^2_+$. See Figure 1.

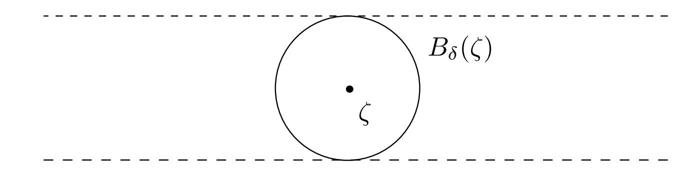


Figure 1. Disc contained in a strip

This gives the following majorization:

$$\begin{aligned}\int_{|z|<\delta} |F(\zeta + z)|^2 \, dx \, dy &\leq \int_{|y|<\delta}\int_{\mathbb{R}} |F(\zeta + x + iy)|^2 \, dx \, dy \\ &\leq 2\delta \sup_{-\delta<y<\delta}\int_{\mathbb{R}} |F(x + i(\eta + y))|^2 \, dx.\end{aligned}$$

Recalling that $\eta > \delta$, we see that the last expression is in fact majorized by

$$2\delta \sup_{y>0}\int_{\mathbb{R}} |F(x + iy)|^2 \, dx = 2\delta \, \|F\|^2_{H^2(\mathbb{R}^2_+)}.$$

In all $|F(\zeta)|^2 \leq \frac{2}{\pi\delta}\|F\|^2_{H^2}$ in the half-plane $\mathrm{Im}(\zeta) > 0$, which proves the lemma.

We now turn to the proof of the identity (7). Starting with F in $H^2(\mathbb{R}^2_+)$, we improve it by replacing it with the function F^ϵ defined by

$$F^\epsilon(z) = F(z)\frac{1}{(1 - i\epsilon z)^2}, \qquad \text{with } \epsilon > 0.$$

Observe that $|F^\epsilon(z)| \leq |F(z)|$ when $\mathrm{Im}(z) > 0$; also $F^\epsilon(z) \to F(z)$ for each such z, as $\epsilon \to 0$. This shows that for each $y > 0$, $F^\epsilon(x + iy) \to$

$F(x+iy)$ in the L^2-norm. Moreover, the lemma guarantees that each F^ϵ satisfies the decay estimate

$$F^\epsilon(z) = O\left(\frac{1}{1+x^2}\right) \qquad \text{whenever } \mathrm{Im}(z) > \delta, \text{ for some } \delta > 0.$$

We assert first that (7) holds with F replaced by F^ϵ. This is a simple consequence of contour integration applied to the function

$$G(z) = F^\epsilon(z)e^{-2\pi i z\xi}.$$

In fact we integrate $G(z)$ over the rectangle with vertices $-R+iy_1$, $R+iy_1$, $R+iy_2$, $-R+iy_2$, and let $R \to \infty$. If we take into account that $G(z) = O(1/(1+x^2))$ in this rectangle, then we find that

$$\int_{L_1} G(z)\,dz = \int_{L_2} G(z)\,dz,$$

where L_j is the line $\{x+iy_j : \; x \in \mathbb{R}\}$, $j = 1, 2$. Since

$$\int_{L_j} G(z)\,dz = \int_{\mathbb{R}} F^\epsilon(x+iy_j)e^{-2\pi i(x+iy_j)\xi}\,dx,$$

This means that

$$\hat{F}^\epsilon_{y_1}(\xi)e^{2\pi y_1 \xi} = \hat{F}^\epsilon_{y_2}(\xi)e^{2\pi y_2 \xi}.$$

Since $F^\epsilon(x+iy_j) \to F(x+iy_j)$ in the L^2-norm as $\epsilon \to 0$, we then obtain (7).

The identity we have just proved states that $\hat{F}_y(\xi)e^{2\pi y\xi}$ is independent of y, $y > 0$, and thus there is a function $\hat{F}_0(\xi)$ so that $\hat{F}_y(\xi)e^{2\pi\xi y} = \hat{F}_0(\xi)$; as a result

$$\hat{F}_y(\xi) = \hat{F}_0(\xi)e^{-2\pi\xi y} \qquad \text{for all } y > 0.$$

Therefore by Plancherel's identity

$$\int_{\mathbb{R}} |F(x+iy)|^2\,dx = \int_{\mathbb{R}} |\hat{F}_0(\xi)|^2 e^{-4\pi\xi y}\,d\xi,$$

and hence

$$\sup_{y>0} \int_{\mathbb{R}} |\hat{F}_0(\xi)|^2 e^{-4\pi\xi y}\,d\xi = \|F\|^2_{H^2(\mathbb{R}^2_+)} < \infty.$$

Finally this in turn implies that $\hat{F}_0(\xi) = 0$ for almost every $\xi \in (-\infty, 0)$. For if this were not the case, then for appropriate positive numbers a, b, and c we could have that $|\hat{F}_0(\xi)| \geq a$ for ξ in a set E in $(-\infty, -b)$, with $m(E) \geq c$. This would give $\int |\hat{F}_0(\xi)|^2 e^{-4\pi\xi y}\, d\xi \geq a^2 c e^{4\pi b y}$, which grows indefinitely as $y \to \infty$. The contradiction thus obtained shows that $\hat{F}_0(\xi)$ vanishes almost everywhere when $\xi \in (-\infty, 0)$.

To summarize, for each $y > 0$ the function $\hat{F}_y(\xi)$ equals $\hat{F}_0(\xi)e^{-2\pi\xi y}$, with $\hat{F}_0 \in L^2(0, \infty)$. The Fourier inversion formula then yields the representation (6) for an arbitrary element of H^2, and the proof of the theorem is concluded.

The second result we deal with may be viewed as the half-plane analogue of Fatou's theorem in the previous chapter.

Theorem 2.3 *Suppose F belongs to $H^2(\mathbb{R}^2_+)$. Then $\lim_{y\to 0} F(x+iy) = F_0(x)$ exists in the following two senses:*

(i) *As a limit in the $L^2(\mathbb{R})$-norm.*

(ii) *As a limit for almost every x.*

Thus F has boundary values (denoted by F_0) in either of the two senses above. The function F_0 is sometimes referred to as the **boundary-value function** of f. The proof of (i) is immediate from what we already know. Indeed, if F_0 is the L^2 function whose Fourier transform is $\hat{F}_0$, then

$$\|F(x+iy) - F_0(x)\|^2_{L^2(\mathbb{R})} = \int_0^\infty |\hat{F}_0(\xi)|^2 |e^{-2\pi\xi y} - 1|^2\, dy,$$

and this tends to zero as $y \to 0$ by the dominated convergence theorem.

To prove the almost everywhere convergence, we establish the **Poisson integral representation**

$$\int_{\mathbb{R}} \hat{f}(\xi) e^{-2\pi|\xi|y} e^{2\pi i x\xi}\, d\xi = \int_{\mathbb{R}} f(x-t)\mathcal{P}_y(t)\, dt, \tag{10}$$

with

$$\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{y^2+x^2}$$

the Poisson kernel.[3] This identity holds for every $(x, y) \in \mathbb{R}^2_+$ and any function f in $L^2(\mathbb{R})$. To see this, we begin by noting the following elementary integration formulas:

$$\int_0^\infty e^{2\pi i\xi z}\, d\xi = \frac{i}{2\pi z} \qquad \text{if } \mathrm{Im}(z) > 0, \tag{11}$$

[3]This is the analogue in $\mathbb{R}$ of the identity (3) for the circle, given in Chapter 4.

and

$$\int_{\mathbb{R}} e^{-2\pi|\xi|y} e^{2\pi i\xi x}\, d\xi = \frac{1}{\pi}\frac{y}{y^2+x^2} \qquad \text{if } y > 0. \tag{12}$$

The first is an immediate consequence of the fact that

$$\int_0^N e^{2\pi i\xi z}\, d\xi = \frac{1}{2\pi i z}[e^{2\pi i N z} - 1]$$

if we let $N \to \infty$. To prove the second formula, we write the integral as

$$\int_0^\infty e^{-2\pi\xi y} e^{2\pi i\xi x}\, d\xi + \int_0^\infty e^{-2\pi\xi y} e^{-2\pi i\xi x}\, d\xi,$$

which equals

$$\frac{i}{2\pi}\left[\frac{1}{x+iy} + \frac{1}{-x+iy}\right] = \frac{1}{\pi}\frac{y}{y^2+x^2}$$

by (11).

Next we establish (10) when f belongs to (say) the space $\mathcal{S}$. Indeed, for fixed $(x,y) \in \mathbb{R}^2_+$ consider the function $\Phi(t,\xi) = f(t)e^{-2\pi i\xi t}e^{-2\pi|\xi|y}e^{2\pi i\xi x}$ on $\mathbb{R}^2 = \{(\xi,t)\}$. Since $|\Phi(t,\xi)| = |f(t)|e^{-2\pi|\xi|y}$, then (because f is rapidly decreasing) Φ is integrable over $\mathbb{R}^2$. Applying Fubini's theorem yields

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}} \Phi(t,\xi)\, d\xi\right) dt = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} \Phi(t,\xi)\, dt\right) d\xi.$$

The right-hand side obviously gives $\int_{\mathbb{R}} \hat{f}(\xi)e^{-2\pi|\xi|y}e^{2\pi i x\xi}\, d\xi$, while the left-hand side yields $\int_{\mathbb{R}} f(t)\mathcal{P}_y(x-y)\, dt$ in view of (12) above. However, if we use the relation (6) in Chapter 2 we see that

$$\int_{\mathbb{R}} f(t)\mathcal{P}_y(x-y)\, dt = \int_{\mathbb{R}} f(x-t)\mathcal{P}_y(t)\, dt.$$

Thus the Poisson integral representation (10) holds for every $f \in \mathcal{S}$. For a general $f \in L^2(\mathbb{R})$ we consider a sequence $\{f_n\}$ of elements in $\mathcal{S}$, so that $f_n \to f$ (and also $\hat{f}_n \to \hat{f}$) in the L^2-norm. A passage to the limit then yields the formula for f from the corresponding formula for each f_n. Indeed, by the Cauchy-Schwarz inequality we have

$$\left|\int_{\mathbb{R}} [\hat{f}(\xi) - \hat{f}_n(\xi)]e^{-2\pi|\xi|y}e^{2\pi i x\xi}\, d\xi\right| \le \|\hat{f} - \hat{f}_n\|_{L^2}\left(\int_{\mathbb{R}} e^{-4\pi|\xi|y}\, d\xi\right)^{1/2},$$

and also

$$\left| \int_{\mathbb{R}} [f(x-t) - f_n(x-t)]\mathcal{P}_y(t)\, dt \right| \leq \|f - f_n\|_{L^2} \left(\int_{\mathbb{R}} |\mathcal{P}_y(t)|^2\, dt \right)^{1/2},$$

and the right-hand sides tend to 0 because for each fixed $(x, y) \in \mathbb{R}^2_+$ the functions $e^{-2\pi|\xi|y}$, $\xi \in \mathbb{R}$, and $\mathcal{P}_y(t)$, $t \in \mathbb{R}$, belong to $L^2(\mathbb{R})$.

Having established the Poisson integral representation (10), we return to our given element $F \in H^2(\mathbb{R}^2_+)$. We know that there is an L^2 function $\hat{F}_0(\xi)$ (which vanishes when $\xi < 0$) such that (6) holds. With F_0 the $L^2(\mathbb{R})$ function whose Fourier transform is $\hat{F}_0(\xi)$, we see from (10), with $f = F_0$, that

$$F(x + iy) = \int_{\mathbb{R}} F_0(x - t)\mathcal{P}_y(t)\, dt.$$

From this we deduce the fact that $F(x + iy) \to F_0(x)$ a.e in x as $y \to 0$, since the family $\{\mathcal{P}_y\}$ is an approximation of the identity for which Theorem 2.1 in Chapter 3 applies. There is, however, one small obstacle that has to be overcome: the theorem as stated applied to L^1 functions and not to functions in L^2. Nevertheless, given the nature of the approximation to the identity, a simple "localization" argument will succeed. We proceed as follows.

It will suffice to see that for any large N, which is fixed, $F(x + iy) \to F_0(x)$, for a.e x with $|x| < N$. To do this, decompose F_0 as $G + H$, where $G(x) = F_0(x)$ when $|x| > 2N$, $G(x) = 0$ when $|x| \geq 2N$; thus $H(x) = 0$ if $|x| \leq 2N$ but $|H(x)| \leq |F_0(x)|$. Note that now $G \in L^1$ and

$$\int_{\mathbb{R}} F_0(x - t)\mathcal{P}_y(t)\, dt = \int_{\mathbb{R}} G(x - t)\mathcal{P}_y(t)\, dt + \int_{\mathbb{R}} H(x - t)\mathcal{P}_y(t)\, dt.$$

Therefore, by the above mentioned theorem in Chapter 3, the first integral on the right-hand side converges for a.e x to $G(x) = F_0(x)$ when $|x| < N$. While when $|x| < N$ the integrand of the second integral vanishes when $|t| < N$ (since then $|x - t| < 2N$). That integral is therefore majorized by

$$\left(\int_{\mathbb{R}} |H(x - t)|^2\, dt \right)^{1/2} \left(\int_{|t| \geq N} |\mathcal{P}_y(t)|^2\, dt \right)^{1/2}.$$

However $\left(\int_{\mathbb{R}} |H(x - t)|^2\, dt \right)^{1/2} \leq \|F_0\|_{L^2}$, while (as is easily seen) $\int_{|t| \geq N} |\mathcal{P}_y(t)|^2\, dt \to 0$ as $y \to 0$. Hence $F(x + iy) \to F_0(x)$ for a.e x with

$|x| < N$, as $y \to 0$, and since N is arbitrary, the proof of Theorem 2.3 is now complete.

The following comments may help clarify the thrust of the above theorems.

(i) Let S be the subspace of $L^2(\mathbb{R})$ consisting of all functions F_0 arising in Theorem 2.3. Then, since the functions F_0 are exactly those functions in L^2 whose Fourier transform is supported on the half-line $(0, \infty)$, we see that S is a closed subspace. We might be tempted to say that S consists of those functions in L^2 that arise as boundary values of holomorphic functions in the upper half-plane; but this heuristic assertion is not exact if we do not add a quantitative restriction such as in the definition (5) of the Hardy space. See Exercise 4.

(ii) Suppose we defined P to be the orthogonal projection on the subspace S of L^2. Then, as is easily seen, $\widehat{(Pf)}(\xi) = \chi(\xi)\hat{f}(\xi)$ for any $f \in L^2(\mathbb{R})$; here χ is the characteristic function of $(0, \infty)$. The operator P is also closely related to the **Cauchy integral**. Indeed, if F is the (unique) element in $H^2(\mathbb{R}^2_+)$ whose boundary function (according to Theorem 2.3) is $P(f)$, then

$$F(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(t)}{t - z}\, dt, \qquad z \in \mathbb{R}^2_+.$$

To prove this it suffices to verify that for any $f \in L^2(\mathbb{R})$ and any fixed $z = x + iy \in \mathbb{R}^2_+$, we have

$$\int_0^\infty \hat{f}(\xi) e^{2\pi i \xi z}\, d\xi = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(t)}{t - z}\, dt.$$

This is proved in the same way as the Poisson integral representation (10) except here we use the identity (11) instead of (12). The details may be left to the interested reader. Also, the reader might note the close analogy between this version of the Cauchy integral for the upper-half plane, and a corresponding version for the unit disc, as given in Example 2, Section 4 of Chapter 4.

(iii) In analogy with the periodic case discussed in Exercise 30 of Chapter 4, we define a **Fourier multiplier operator** T on $\mathbb{R}$ to be a linear operator on $L^2(\mathbb{R})$ determined by a bounded function m (the **multiplier**), such that T is defined by the formula $\widehat{(Tf)}(\xi) = m(\xi)\hat{f}(\xi)$ for any $f \in L^2(\mathbb{R})$. The orthogonal projection P above is such an operator and its multiplier is the characteristic function $\chi(\xi)$. Another closely related operator of this type is the **Hilbert transform** H defined by

$P = \frac{I+iH}{2}$. Then H is a Fourier multiplier operator corresponding to the multiplier $\frac{1}{i}\text{sign}(\xi)$. Among the many important properties of H is its connection to conjugate harmonic functions. Indeed, for f a real-valued function in $L^2(\mathbb{R})$, f and $H(f)$ are, respectively, the real and imaginary parts of the boundary values of a function in the Hardy space. More about the Hilbert transform can be found in Exercises 9 and 10 and Problem 5 below.

3 Constant coefficient partial differential equations

We turn our attention to solving the linear partial differential equation

$$L(u) = f, \tag{13}$$

where the operator L takes the form

$$L = \sum_{|\alpha| \leq n} a_\alpha \left(\frac{\partial}{\partial x}\right)^\alpha$$

with $a_\alpha \in \mathbb{C}$ constants.

In the study of the classical examples of L, such as the wave equation, the heat equation, and Laplace's equation, one already sees the Fourier transform entering in an important way.[4] For general L, this key role is further indicated by the following simple observation. If, for example, we try to solve this equation with both u and f elements in $\mathcal{S}$, then this is equivalent to the algebraic equation

$$P(\xi)\hat{u}(\xi) = \hat{f}(\xi),$$

where $P(\xi)$ is the **characteristic polynomial** of f defined by

$$P(\xi) = \sum_{|\alpha| \leq n} a_\alpha (2\pi i \xi)^\alpha.$$

This is because one has the Fourier transform identity

$$\widehat{\left(\frac{\partial^\alpha f}{\partial x^\alpha}\right)}(\xi) = (2\pi i\xi)^\alpha \hat{f}(\xi).$$

Thus a solution u in the space $\mathcal{S}$ (if it exists) would be uniquely determined by

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{P(\xi)}.$$

[4]See for example Chapters 5 and 6 in Book I.

In a more general setting, matters are not so easy: aside from the question of defining (13), the Fourier transform is not directly applicable; also, solutions that we prove to exist (but are not unique!) have to be understood in a wider sense.

3.1 Weak solutions

As the reader may have guessed, it will not suffice to restrict our attention to those functions for which $L(u)$ is defined in the usual way, but instead a broader notion is needed, one involving the idea of "weak solutions." To describe this concept, we start with a given open set Ω in $\mathbb{R}^d$ and consider the space $C_0^\infty(\Omega)$, which consists of the indefinitely differentiable functions[5] having compact support in Ω.[6] We have the following fact.

Lemma 3.1 *The space $C_0^\infty(\Omega)$ is dense in $L^2(\Omega)$ in the norm $\|\cdot\|_{L^2(\Omega)}$.*

The proof is essentially a repetition of that of Lemma 1.2. We take the precaution of modifying the definition of g_M given there to be: $g_M(x) = f(x)$ if $|x| \leq M$, $d(x, \Omega^c) \geq 1/M$ and $|f(x)| \leq M$, and $g_M(x) = 0$ otherwise. Also, when we regularize g_M, we replace it with $g_M * \varphi_\delta$, with $\delta < 1/2M$. Then the support of $g_M * \varphi_\delta$ is still compact and at a distance $\geq 1/2M$ from Ω^c.

We next consider the **adjoint operator** of L defined by

$$L^* = \sum_{|\alpha| \leq n} (-1)^{|\alpha|} \overline{a_\alpha} \left(\frac{\partial}{\partial x} \right)^\alpha .$$

The operator L^* is called the adjoint of L because, in analogy with the definition of the adjoint of a bounded linear transformation given in Section 5.2 of the previous chapter, we have

$$(14) \qquad (L\varphi, \psi) = (\varphi, L^*\psi) \quad \text{whenever } \varphi, \psi \in C_0^\infty(\Omega),$$

where $(\cdot, \cdot)$ denotes the inner product on $L^2(\Omega)$ (which is the restriction of the usual inner product on $L^2(\mathbb{R}^d)$). The identity (14) is proved by successive integration by parts. Indeed, consider first the special case when $L = \partial/\partial x_j$, and then $L^* = -\partial/\partial x_j$. If we use Fubini's theorem, integrating first in the x_j variable, then in this case (14) reduces to the

[5]Indefinitely differentiable functions are also referred to as C^∞ functions, or **smooth** functions.

[6]This means that the closure of the support of f, as defined in Section 1 of Chapter 2, is compact and contained in Ω.

familiar one-dimensional formula

$$\int_{-\infty}^{\infty} \left(\frac{d\varphi}{dx}\right) \overline{\psi}\, dx = -\int_{-\infty}^{\infty} \varphi \left(\frac{d\overline{\psi}}{dx}\right) dx,$$

with the integrated boundary terms vanishing because of the assumed support properties of ψ (or φ). Once established for $L = \partial/\partial x_j$, $1 \leq j \leq n$, then (14) follows for $L = (\partial/\partial x)^\alpha$ by iteration, and hence for general L by linearity.

At this point we digress momentarily to consider besides $C_0^\infty(\Omega)$ some other spaces of differentiable functions on Ω that will be useful later. The space $C^n(\Omega)$ consists of all functions f on Ω that have continuous partial derivatives of order $\leq n$. Also, the space $C^n(\overline{\Omega})$ consists of those functions on $\overline{\Omega}$ that can be extended to functions in $\mathbb{R}^d$ that belong to $C^n(\mathbb{R}^d)$. Thus, in an obvious sense, we have the inclusion relation

$$C_0^\infty(\Omega) \subset C^n(\overline{\Omega}) \subset C^n(\Omega), \quad \text{for each positive integer } n.$$

Returning to our partial differential operator L, it is useful to observe that the formula

$$(Lu, \psi) = (u, L^*\psi)$$

continues to hold (with the same proof) if we merely assume that $u \in C^n(\Omega)$ without assuming it has compact support, while still supposing $\psi \in C_0^\infty(\Omega)$.

In particular, if we have $L(u) = f$ in the ordinary sense (sometimes called the "strong" sense), which requires the assumption that $u \in C^n(\Omega)$ in order to define the partial derivatives entering in Lu, then we would also have

$$(f, \psi) = (u, L^*\psi) \quad \text{for all } \psi \in C_0^\infty(\Omega). \tag{15}$$

This leads to the following important definition: if $f \in L^2(\Omega)$, a function $u \in L^2(\Omega)$ is a **weak solution** of the equation $Lu = f$ in Ω if (15) holds. Of course an ordinary solution is always a weak solution.

Significant instances of weak solutions that are not ordinary solutions already arise in elementary situations such as in the study of the one-dimensional wave equation. Here $L(u) = (\partial^2 u/\partial x^2) - (\partial^2 u/\partial t^2)$, so the underlying space is $\mathbb{R}^2 = \{(x_1, x_2) : \text{with } x_1 = x,\ x_2 = t\}$. Suppose, for example, we consider the case of the "plucked string."[7] We are then

[7]See Chapter 1 in Book I.

looking at the solution of $L(u) = 0$ subject to the boundary conditions $u(x, 0) = f(x)$ and $(\partial u/\partial t)(x, 0) = 0$ for $0 \leq x \leq \pi$, where the graph of f is piecewise linear and is illustrated in Figure 2.

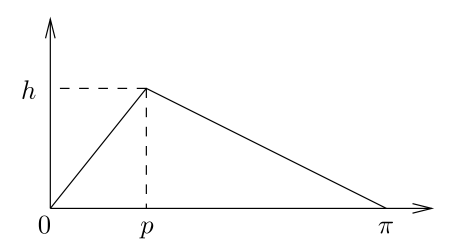


Figure 2. Initial position of a plucked string

If one extends f to $[-\pi, \pi]$ by making it odd, and then to all of $\mathbb{R}$ by periodicity (of period 2π), then the solution is given by d'Alembert's formula

$$u(x,t) = \frac{f(x+t) + f(x-t)}{2}.$$

In the present case u is *not* twice continuously differentiable, and it is therefore not an ordinary solution. Nevertheless it is a weak solution. To see this, approximate f by a sequence of functions f_n that are C^∞ and such that $f_n \to f$ uniformly on every compact subset of $\mathbb{R}$.[8] If we define $u_n(x,t)$ as $[f_n(x+t) + f_n(x-t)]/2$, we can check directly that $L(u_n) = 0$ and hence $(u_n, L^*\psi) = 0$ for all $\psi \in C_0^\infty(\mathbb{R}^2)$, and thus by uniform convergence we obtain that $(u, L^*\psi) = 0$ as desired.

A different example illustrating the nature of weak solutions arises for the operator $L = d/dx$ on $\mathbb{R}$. If we suppose $\Omega = (0,1)$, then with u and f in $L^2(\Omega)$, we have that $Lu = f$ in the weak sense if and only if there is an absolutely continuous function F on $[0,1]$ such that $F(x) = u(x)$ and $F'(x) = f(x)$ almost everywhere. For more about this, see Exercise 14.

3.2 The main theorem and key estimate

We now turn to the general theorem guaranteeing the existence of solutions of partial differential equations with constant coefficients

Theorem 3.2 *Suppose Ω is a bounded open subset of $\mathbb{R}^d$. Given a linear partial differential operator L with constant coefficients, there exists a*

[8]One may write, for example, $f_n = f * \varphi_{1/n}$, where $\{\varphi_\epsilon\}$ is the approximation to the identity, as in the proof of Lemma 1.2.

bounded linear operator K on $L^2(\Omega)$ such that whenever $f \in L^2(\Omega)$, then

$$L(Kf) = f \quad \textit{in the weak sense.}$$

In other words, $u = K(f)$ is a weak solution to $L(u) = f$.

The heart of the matter lies in an inequality that we state next, but whose proof (which uses the Fourier transform) is postponed until the next section.

Lemma 3.3 *There exists a constant c such that*

$$\|\psi\|_{L^2(\Omega)} \leq c\|L^*\psi\|_{L^2(\Omega)} \quad \textit{whenever } \psi \in C_0^\infty(\Omega).$$

The usefulness of this lemma comes about for the following reason. If L is a finite-dimensional linear transformation, the solvability of L (the fact that it is surjective) is of course equivalent with the fact that its adjoint L^* is injective. In effect, the lemma provides the analytic substitute for this reasoning in an infinite-dimensional setting.

We first prove the theorem assuming the validity of the inequality in the lemma.

Consider the pre-Hilbert space $\mathcal{H}_0 = C_0^\infty(\Omega)$ equipped with the inner product and norm

$$\langle \varphi, \psi \rangle = (L^*\varphi, L^*\psi), \qquad \|\psi\|_0^2 = \|L^*\psi\|_{L^2(\Omega)}.$$

Following the results in Section 2.3 of Chapter 4, we let $\mathcal{H}$ denote the completion of $\mathcal{H}_0$. By Lemma 3.3, a Cauchy sequence in the $\|\cdot\|_0$-norm is also Cauchy in the $L^2(\Omega)$-norm; hence we may identify $\mathcal{H}$ with a subspace of $L^2(\Omega)$. Also, L^*, initially defined as a bounded operator from $\mathcal{H}_0$ to $L^2(\Omega)$, extends to a bounded operator L^* from $\mathcal{H}$ to $L^2(\Omega)$ (by Lemma 1.3). For a fixed $f \in L^2(\Omega)$, consider the linear map $\ell_0 : C_0^\infty(\Omega) \to \mathbb{C}$ defined by

$$\ell_0(\psi) = (\psi, f) \quad \text{for } \psi \in C_0^\infty(\Omega).$$

The Cauchy-Schwarz inequality together with another application of Lemma 3.3 yields

$$\begin{aligned} |\ell_0(\psi)| = |(\psi, f)| &\leq \|\psi\|_{L^2(\Omega)}\|f\|_{L^2(\Omega)} \\ &\leq c\|L^*\psi\|_{L^2(\Omega)}\|f\|_{L^2(\Omega)} \\ &\leq c'\|\psi\|_0, \end{aligned}$$

with $c' = c\|f\|_{L^2(\Omega)}$. Hence ℓ_0 is bounded on the pre-Hilbert space $\mathcal{H}_0$. Therefore, ℓ extends to a bounded linear functional on $\mathcal{H}$ (see Section 5.1, Chapter 4), and the above inequalities show that $\|\ell\| \leq c\|f\|_{L^2(\Omega)}$. By the Riesz representation theorem applied to ℓ on the Hilbert space $\mathcal{H}$ (Theorem 5.3 in Chapter 4), there exists $U \in \mathcal{H}$ such that

$$\ell(\psi) = \langle \psi, U \rangle = (L^*\psi, L^*U) \quad \text{ for all } \psi \in C_0^\infty(\Omega).$$

Here $\langle \cdot, \cdot \rangle$ denotes the extension to $\mathcal{H}$ of the initial inner product on $\mathcal{H}_0$, and L^* also denotes the extension of L^* originally given on $\mathcal{H}_0$.

If we let $u = L^*U$, then $u \in L^2(\Omega)$, and we find that

$$\ell(\psi) = (\psi, f) = (L^*\psi, u) \quad \text{ for all } \psi \in C_0^\infty(\mathbb{R}^d).$$

Hence

$$(f, \psi) = (u, L^*\psi) \quad \text{ for all } \psi \in C_0^\infty(\mathbb{R}^d),$$

and by definition, u is a weak solution to the equation $Lu = f$ in Ω. If we let $Kf = u$, we see that once f is given, Kf is uniquely determined by the above steps. Since $\|U\|_0 = \|\ell\| \leq c\|f\|_{L^2(\Omega)}$ we see that

$$\|Kf\|_{L^2(\Omega)} = \|u\|_{L^2(\Omega)} = \|L^*U\|_{L^2(\Omega)} = \|U\|_0 \leq c\|f\|_{L^2(\Omega)},$$

whence $K : L^2(\Omega) \to L^2(\Omega)$ is bounded.

Proof of the main estimate

To complete the proof of the theorem, we must still prove the estimate in Lemma 3.3, that is,

$$\|\psi\|_{L^2(\Omega)} \leq c\|L^*\psi\|_{L^2(\Omega)} \quad \text{ whenever } \psi \in C_0^\infty(\Omega).$$

The reasoning below relies on an important fact: if f has compact support in $\mathbb{R}$, then $\hat{f}(\xi)$ initially defined for $\xi \in \mathbb{R}$ extends to an entire function for $\zeta = \xi + i\eta \in \mathbb{C}$. This observation reduces the problem to an inequality about holomorphic functions and polynomials.

Lemma 3.4 *Suppose $P(z) = z^m + \cdots + a_1 z + a_0$ is a polynonial of degree m with leading coefficient 1. If F is a holomorphic function on $\mathbb{C}$, then*

$$|F(0)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |P(e^{i\theta})F(e^{i\theta})|^2 d\theta.$$

Proof. The lemma is a consequence of the special case when $P = 1$

$$|F(0)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} |F(e^{i\theta})|^2 \, d\theta. \tag{16}$$

This assertion follows directly from the mean-value identity (8) in Section 2 with $\zeta = 0$ and $r = 1$, via the Cauchy-Schwarz inequality. With it we begin by factoring P:

$$P(z) = \prod_{|\alpha| \geq 1} (z - \alpha) \prod_{|\beta| < 1} (z - \beta) = P_1(z) P_2(z),$$

where each product is finite and taken over the roots of P whose absolute values are ≥ 1 and < 1, respectively.

Note that $|P_1(0)| = \prod_{|\alpha| \geq 1} |\alpha| \geq 1$.

For P_2 we write

$$(z - \beta) = -(1 - \overline{\beta} z) \psi_\beta(z),$$

where $\psi_\beta(z) = \frac{\beta - z}{1 - \overline{\beta} z}$ are the "Blaschke factors" that have the obvious property that they are holomorphic in a region containing the closed unit disc and $|\psi_\beta(e^{i\theta})| = 1$; see also Chapter 8 in Book II. We write $\tilde{P}_2 = \prod_{|\beta| < 1} (1 - \overline{\beta} z)$ and $\tilde{P} = P_1 \tilde{P}_2$. Thus $|\tilde{P}(0)| \geq 1$, while $|\tilde{P}(e^{i\theta})| = |P(e^{i\theta})|$ for every θ. We now apply (16) to the function $\tilde{P} F$ in place of F and find that

$$\begin{aligned} |F(0)|^2 \leq |\tilde{P}(0) F(0)|^2 &\leq \frac{1}{2\pi} \int_0^{2\pi} |\tilde{P}(e^{i\theta}) F(e^{i\theta})|^2 \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} |P(e^{i\theta}) F(e^{i\theta})|^2 \, d\theta, \end{aligned}$$

which gives the desired conclusion.

We turn to the proof of the inequality $\|\psi\| \leq c \|L^* \psi\|$ for all $\psi \in C_0^\infty(\Omega)$ in the special case of one dimension, that is, $\Omega \subset \mathbb{R}$.

Suppose f is an L^2 function supported on the interval $[-M, M]$. Then

$$\hat{f}(\xi) = \int_{-M}^{M} f(x) e^{-2\pi i x \xi} \, dx$$

whenever $\xi \in \mathbb{R}$. In fact, the above integral converges whenever ξ is replaced by $\zeta = \xi + i\eta \in \mathbb{C}$, and we may extend $\hat{f}$ to a holomorphic function of ζ in the whole complex plane. An application of the Plancherel

formula (for fixed η) yields

$$\int_{-\infty}^{\infty} |\hat{f}(\xi + i\eta)|^2 \, d\xi \leq e^{4\pi M|\eta|} \int_{-\infty}^{\infty} |f(x)|^2 \, dx.$$

We use this observation in the following context. We may assume (upon multiplying L by a suitable constant) that

$$L^* = \sum_{0 \leq k \leq n} (-1)^k \overline{a_k} \left(\frac{\partial}{\partial x}\right)^k,$$

where $a_n = (2\pi i)^{-n}$. If we let $Q(\xi) = \sum_{0 \leq k \leq n} (-1)^k \overline{a_k} (2\pi i \xi)^k$ be its characteristic polynomial, then we note that

$$\widehat{L^*\psi}(\xi) = Q(\xi)\hat{\psi}(\xi) \qquad \text{whenever } \psi \in C_0^\infty(\mathbb{R}).$$

If M is chosen so large that $\Omega \subset [-M, M]$, then our previous observation gives

$$\int_{-\infty}^{\infty} |Q(\xi + i\eta)\hat{\psi}(\xi + i\eta)|^2 d\xi \leq e^{4\pi M|\eta|} \int_{-\infty}^{\infty} |L^*\psi(x)|^2 dx. \tag{17}$$

Picking $\eta = i \sin\theta$, and making a translation by $\cos\theta$ yields

$$\int_{-\infty}^{\infty} |Q(\xi + \cos\theta + i\sin\theta)\hat{\psi}(\xi + \cos\theta + i\sin\theta)|^2 d\xi \leq$$

$$\leq e^{4\pi M} \int_{-\infty}^{\infty} |L^*\psi(x)|^2 dx.$$

An application of Lemma 3.4 with $F(z) = \hat{\psi}(\xi + z)$ and $Q(\xi + z)$ in place of $P(z)$ then gives

$$|\hat{\psi}(\xi)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |Q(\xi + \cos\theta + i\sin\theta)\hat{\psi}(\xi + \cos\theta + i\sin\theta)|^2 d\theta.$$

We now integrate in ξ over $\mathbb{R}$, and on the right-hand side interchange the order of the ξ and θ integrations; also by translation invariance we replace the integration in the ξ variable by that in the variable $\xi + \cos\theta$. Using (17) the result is

$$\begin{aligned}\|\hat{\psi}\|_{L^2(\mathbb{R})}^2 &\leq \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} |Q(\xi + i\sin\theta)\hat{\psi}(\xi + i\sin\theta)|^2 \, d\xi \, d\theta \\ &\leq e^{4\pi M} \int_{\mathbb{R}} |L^*\psi(x)|^2 \, dx,\end{aligned}$$

which by Plancherel's identity proves the main lemma in the one-dimensional case.

The higher dimensional case is a modification of the argument above. Let $Q = \sum_{|\alpha|\leq n}(-1)^{\alpha}a_{\alpha}(2\pi i\xi)^{\alpha}$ be the characteristic polynomial of L^*. Then we can choose a new set of orthogonal axes (whose coordinates we denote by $(\xi_1, \ldots, \xi_d)$) so that if $\xi = (\xi_1, \xi')$ with $\xi' = (\xi_2, \ldots, \xi_d)$, then after multiplying by a suitable constant

$$Q(\xi) = (2\pi i)^{-n}\xi_1^n + \sum_{j=0}^{n-1} \xi_1^j q_j(\xi'), \tag{18}$$

where $q_j(\xi')$ are polynomials of ξ' (of degrees $\leq n - j$).

To see that such a choice is possible, write $Q = Q_n + Q'$, where Q_n is homogeneous of degree n and Q' has degree $< n$. Then since we may assume $Q_n \neq 0$ there is (after multiplying Q by a suitable constant), a unit vector γ so that $Q_n(\gamma) = (2\pi i)^{-n}$. Then $Q_n(\xi) = (2\pi i)^{-n}r^n$ if $\xi = \gamma r$, $r \in \mathbb{R}$. We can then take the ξ_1-axis to lie along γ, and the $\xi_2, \ldots, \xi_d$-axes to be in mutually orthogonal directions, from which the form (18) is clear.

Proceeding now as before we obtain

$$|\hat{\psi}(\xi_1, \xi')|^2 \leq \frac{1}{2\pi}\int_0^{2\pi} |Q(\xi_1 + e^{i\theta}, \xi')\hat{\psi}(\xi_1 + e^{i\theta}, \xi')|^2\, d\theta$$

for each $(\xi_1, \xi') \in \mathbb{R}^d$. An integration[9] then gives

$$\|\hat{\psi}\|^2_{L^2(\mathbb{R}^d)} \leq \frac{1}{2\pi}\int_0^{2\pi}\int_{\mathbb{R}^d} |Q(\xi_1 + i\sin\theta, \xi')\hat{\psi}(\xi_1 + i\sin\theta, \xi')|^2\, d\xi\, d\theta.$$

If we suppose that the projection of the (bounded) set Ω on the x_1-axis is contained in $[-M, M]$, we see as before that the right-hand side above is majorized by $e^{4\pi M}\int_{\mathbb{R}^d} |L^*\psi(x)|^2\, dx$, finishing the proof of Lemma 3.3 and hence that of the theorem.

4* The Dirichlet principle

Dirichlet's principle arose in the study of the boundary-value problem for Laplace's equation. Stated in the case of two dimensions it refers to the classical problem of finding the steady-state temperature of a plate

[9]We note that by the rotational invariance of Lebesgue measure (Problem 4 in Chapter 2 and Exercise 26 in Chapter 3), integration in ξ can be carried out in the new coordinates as well.

whose boundary is exposed to a given temperature distribution. The issue raised is the following question, called the **Dirichlet problem**: If Ω is a bounded open set in $\mathbb{R}^2$ and f a continuous function on the boundary $\partial\Omega$, we wish to find a function $u(x_1, x_2)$ such that

$$\begin{cases} \triangle u = 0 & \text{in } \Omega, \\ u = f & \text{on } \partial\Omega. \end{cases} \tag{19}$$

Thus we need to determine a function that is C^2 (twice continuously differentiable) in Ω, whose Laplacian[10] is zero, and which is continuous on the closure of Ω, with $u|_{\partial\Omega} = f$.

With either Ω or f satisfying special symmetry conditions, the solution to this problem can sometimes be written out explicitly. For instance, if Ω is the unit disc, then

$$u(re^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\varphi)P_r(\theta - \varphi)\,d\varphi,$$

where P_r is the Poisson kernel (for the disc). We also obtained (in Books I and II) explicit formulas for the solution of the Dirichlet problem for some unbounded domains. For example, when Ω is the upper half-plane the solution is

$$u(x, y) = \int_{\mathbb{R}} \mathcal{P}_y(x - t)f(t)\,dt,$$

where $\mathcal{P}_y(x)$ is the analogous Poisson kernel for the upper half-plane. A somewhat similar convolution formula was obtained when Ω is a strip. Also, the Dirichlet problem can be solved explicitly for certain Ω by using conformal mappings.[11]

In general, however, there are no explicit solutions, and other methods must be found. An idea that was used intially was based on an approach of wide utility in mathematics and physics: to find the equilibrium state of a system one seeks to minimize an appropriate "energy" or "action." In the present case the role of this energy is played by the **Dirichlet integral**, which is defined for appropriate functions U by

$$\mathcal{D}(U) = \int_{\Omega} |\nabla U|^2 = \int_{\Omega} \left|\frac{\partial U}{\partial x_1}\right|^2 + \left|\frac{\partial U}{\partial x_2}\right|^2 dx_1 dx_2.$$

(Note the similarity with the expression of the "potential energy" in the case of the vibrating string in Chapters 3 and 6 of Book I.) In fact,

[10]The **Laplacian** of a function u in $\mathbb{R}^d$ is defined by $\triangle u = \sum_{k=1}^{d} \partial^2 u/\partial x_k^2$.

[11]The close relation between conformal maps and the Dirichlet problem is discussed in the last part of Section 1 of Chapter 8, in Book II.

that approach underlies the proof Riemann proposed for his well-known mapping theorem. About this early history R. Courant has written:

> Already some years before the rise of Riemann's genius, C.F. Gauss and W. Thompson had observed that the boundary value problem of the harmonic differential equation $\triangle u = u_{xx} + u_{yy} = 0$ for a domain G in the x, y-plane can be reduced to the problem of minimizing the integral $\mathcal{D}[\phi]$ for the domain G, under the condition that the functions ϕ admitted to competition have the prescribed boundary values. Because of the positive character of $\mathcal{D}[\phi]$ the existence of a solution for the latter problem was considered obvious and hence the existence for the former assured. As a student in Dirichlet's lectures, Riemann had been fascinated by this convincing argument: soon afterwards he used it, under the name "Dirichlet's Principle," in a more varied and spectacular manner as the very foundation of his new geometric function theory.

The application of Dirichlet's principle was thought to have been justified by the following simple observation:

Proposition 4.1 *Suppose there exists a function $u \in C^2(\overline{\Omega})$ that minimizes $\mathcal{D}(U)$ among all $U \in C^2(\overline{\Omega})$ with $U|_{\partial\Omega} = f$. Then u is harmonic in Ω.*

Proof. For functions F and G in $C^2(\overline{\Omega})$ define the following inner-product

$$\langle F, G\rangle = \int_\Omega \left(\frac{\partial F}{\partial x_1}\overline{\frac{\partial G}{\partial x_1}} + \frac{\partial F}{\partial x_2}\overline{\frac{\partial G}{\partial x_2}} \right) dx_1 dx_2.$$

We then note that $\mathcal{D}(u) = \langle u, u\rangle$. If v is any function in $C^2(\overline{\Omega})$ with $v|_{\partial\Omega} = 0$, then for all ϵ we have

$$\mathcal{D}(u + \epsilon v) \geq \mathcal{D}(u),$$

since $u + \epsilon v$ and u have the same boundary values, and u minimizes the Dirichlet integral. We note, however, that

$$\mathcal{D}(u + \epsilon v) = \mathcal{D}(u) + \epsilon^2 \mathcal{D}(v) + \epsilon\langle u, v\rangle + \epsilon\langle v, u\rangle.$$

Hence

$$\epsilon^2 \mathcal{D}(v) + \epsilon\langle u, v\rangle + \epsilon\langle v, u\rangle \geq 0,$$

and since ϵ can be both positive or negative, this can happen only if $\text{Re}\langle u, v\rangle = 0$. Similarly, considering the perturbation $u + i\epsilon v$, we find $\text{Im}\langle u, v\rangle = 0$, and therefore $\langle u, v\rangle = 0$. An integration by parts then provides

$$0 = \langle u, v\rangle = -\int_{\Omega} (\triangle u)\overline{v}$$

for all $v \in C^2(\overline{\Omega})$ with $v|_{\partial\Omega} = 0$. This implies that $\triangle u = 0$ in Ω, and of course u equals f on the boundary.

Nevertheless, several serious objections were later raised to Dirichlet's principle. The first was by Weierstrass, who pointed out that it was not clear (and had not been proved) that a minimizing function for the Dirichlet integral exists, so there might simply be no winner to the implied competition in Proposition 4.1. He argued by analogy with a simpler one-dimensional problem: that of minimizing the integral

$$D(\varphi) = \int_{-1}^{1} |x\varphi'(x)|^2 dx$$

among all C^1 functions on $[-1, 1]$ that satisfy $\varphi(-1) = -1$ and $\varphi(1) = 1$. The minimum value achieved by this integral is zero. To verify this, let ψ be a smooth non-decreasing function on $\mathbb{R}$ that satisfies $\psi(x) = 1$ for $x \geq 1$, and $\psi(x) = -1$ if $x \leq -1$. For each $0 < \epsilon < 1$, we consider the function

$$\varphi_\epsilon(x) = \begin{cases} 1 & \text{if } \epsilon \leq x, \\ \psi(x/\epsilon) & \text{if } -\epsilon < x < \epsilon, \\ -1 & \text{if } x \leq -\epsilon. \end{cases}$$

Then φ_ϵ satisfies the desired constraints, and if M denotes a bound for the derivative of ψ, we find

$$\begin{aligned} D(\varphi_\epsilon) &= \int_{-\epsilon}^{\epsilon} |x|^2 |\epsilon^{-1}\psi'(x/\epsilon)|^2 dx \\ &\leq \int_{-\epsilon}^{\epsilon} |\psi'(x/\epsilon)|^2 dx \\ &\leq 2\epsilon M^2. \end{aligned}$$

In the limit as ϵ tends to 0, we find that the minimum value of the integral $D(\varphi)$ is zero. This minimum value cannot be reached by a C^1 function satisfying the boundary conditions, since $D(\varphi) = 0$ implies $\varphi'(x) = 0$ and thus φ is constant.

A further objection was raised by Hadamard, who remarked that $\mathcal{D}(u)$ may be infinite even for a solution u of the boundary value problem: thus, in effect, there may simply be no competitors who qualify for the competition!

To illustrate this point, we return to the disc, and consider the function

$$f(\theta) = f_\alpha(\theta) = \sum_{n=0}^{\infty} 2^{-n\alpha} e^{i2^n\theta}$$

for $\alpha > 0$. This function first appeared in Chapter 4 of Book I, where it is shown that f_α is continuous but nowhere differentiable if $\alpha \leq 1$. The solution of the Dirichlet problem on the unit disc with boundary value f_α is given by the Poisson integral

$$u(r,\theta) = \sum_{n=0}^{\infty} r^{2^n} 2^{-n\alpha} e^{i2^n\theta}.$$

However, the use of polar coordinates gives

$$\left|\frac{\partial u}{\partial x_1}\right|^2 + \left|\frac{\partial u}{\partial x_2}\right|^2 = \left|\frac{\partial u}{\partial r}\right|^2 + \frac{1}{r^2}\left|\frac{\partial u}{\partial \theta}\right|^2.$$

Thus

$$\iint_{D_\rho} \left(\left|\frac{\partial u}{\partial x_1}\right|^2 + \left|\frac{\partial u}{\partial x_2}\right|^2\right) dx_1 dx_2 = \int_0^\rho \int_0^{2\pi} \left(\left|\frac{\partial u}{\partial r}\right|^2 + \frac{1}{r^2}\left|\frac{\partial u}{\partial \theta}\right|^2\right) d\theta r dr$$

where D_ρ is the disc of radius $0 < \rho < 1$ centered at the origin. Since

$$\frac{\partial u}{\partial r} \sim \sum 2^n 2^{-n\alpha} r^{2^n - 1} e^{i2^n\theta} \quad \text{and} \quad \frac{\partial u}{\partial \theta} \sim \sum r^{2^n} 2^{-n\alpha} i 2^n e^{i2^n\theta},$$

applications of Parseval's identity lead to

$$\begin{aligned}\iint_{D_\rho} \left(\left|\frac{\partial u}{\partial x_1}\right|^2 + \left|\frac{\partial u}{\partial x_2}\right|^2\right) dx_1 dx_2 &\approx \int_0^\rho \sum_{n=0}^{\infty} 2^{2n+1} 2^{-2n\alpha} r^{2^{n+1}-1} dr \\ &= \sum_{n=0}^{\infty} \rho^{2^{n+1}} 2^n 2^{-2n\alpha},\end{aligned}$$

which tends to infinity as $\rho \to 1$ if $\alpha \leq 1/2$.

One can formulate this objection in a more precise way by appealing to the result in Exercise 20.

Despite these significant difficulties, Dirichlet's principle can indeed be validated, if applied in the appropriate way. A key insight is that the space of competing functions arising in the proof of the above proposition is itself a pre-Hilbert space, with inner product $\langle \cdot, \cdot \rangle$ given there. The desired solution lies in the completion of this pre-Hilbert space, and this requires the L^2 theory for its analysis. These ideas were clearly not available at the time Dirichlet's principle was first formulated and used.

In what follows we shall describe how these additional concepts can be exploited. We will begin our presentation in the more general d-dimensional setting, but conclude with the application of these techniques to the solution of the two-dimensional problem (19). As an important preliminary matter we start with the study of some basic properties of harmonic functions.

4.1 Harmonic functions

Throughout this section Ω will denote an open subset of $\mathbb{R}^d$. A function u is **harmonic** in Ω if it is twice continuously differentiable[12] and u solves

$$\triangle u = \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2} = 0.$$

We shall see that harmonic functions can be characterized by a number of equivalent properties.[13] Adapting the terminology used in Section 3, we say that u is **weakly harmonic** in Ω if

$$(u, \triangle\psi) = 0 \quad \text{for every } \psi \in C_0^\infty(\Omega). \tag{20}$$

Note that the left-hand side of (20) is well-defined for any u that is integrable on compact subsets of Ω. Thus, in particular, a weakly harmonic function needs to be defined only almost everywhere. Clearly, however, any harmonic function is weakly harmonic.

Another notion is the **mean-value property** generalizing the identity (9) in Section 2 for holomorphic functions. A continuous function u defined in Ω satisfies this property if

$$u(x_0) = \frac{1}{m(B)} \int_B u(x)\, dx \tag{21}$$

for each ball B whose center is x_0 and whose closure $\overline{B}$ is contained in Ω.

[12]In other words, u is in $C^2(\Omega)$ in the notation of Section 3.1.

[13]Note that in the case of one dimension, harmonic functions are linear and so their theory is essentially trivial.

The following two theorems give alternative characterizations of harmonic functions. Their proofs are closely intertwined.

Theorem 4.2 *If u is harmonic in Ω, then u satisfies the mean-value property* (21). *Conversely, a continuous function satisfying the mean-value property is harmonic.*

Theorem 4.3 *Any weakly harmonic function u in Ω can be corrected on a set of measure zero so that the resulting function is harmonic in Ω.*

The above statement says that for a given weakly harmonic function u there exists a harmonic function $\tilde{u}$, so that $\tilde{u}(x) = u(x)$ for a.e. $x \in \Omega$. Notice since $\tilde{u}$ is necessarily continuous it is uniquely determined by u.

Before we prove the theorems, we deduce a noteworthy corollary. It is a version of the **maximum principle**.

Corollary 4.4 *Suppose Ω is a bounded open set, and let $\partial\Omega = \overline{\Omega} - \Omega$ denote its boundary. Assume that u is continuous in $\overline{\Omega}$ and is harmonic in Ω. Then*

$$\max_{x \in \overline{\Omega}} |u(x)| = \max_{x \in \partial\Omega} |u(x)|.$$

Proof. Since the sets $\overline{\Omega}$ and $\partial\Omega$ are compact and u is continuous, the two maxima above are clearly attained. We suppose that $\max_{x \in \overline{\Omega}} |u(x)|$ is attained at an interior point $x_0 \in \Omega$, for otherwise there is nothing to prove.

Now by the mean-value property, $|u(x_0)| \leq \frac{1}{m(B)} \int_B |u(x)|\, dx$. If for some point $x' \in B$ we had $|u(x')| < |u(x_0)|$, then a similar inequality would hold in a small neighborhood of x', and since $|u(x)| \leq |u(x_0)|$ throughout B, the result would be that $\frac{1}{m(B)} \int_B |u(x)|\, dx < |u(x_0)|$, which is a contradiction. Hence $|u(x)| = |u(x_0)|$ for each $x \in B$. Now this is true for each ball B_r of radius r, centered at x_0, such that $B_r \subset \Omega$. Let r_0 be the least upper bound of such r; then $\overline{B}_{r_0}$ intersects the boundary Ω at some point $\tilde{x}$. Since $|u(x)| = |u(x_0)|$ for all $x \in \overline{B}_r$, $r < r_0$, it follows by continuity that $|u(\tilde{x})| = |u(x_0)|$, proving the corollary.

Turning to the proofs of the theorems, we first establish a variant of Green's formula (for the unit ball) that does not explicitly involve boundary terms.[14] Here u, v, and η are assumed to be twice continuously differentiable functions in a neighborhood of the closure of B, but η is also supposed to be supported in a compact subset of B.

---

[14]The more usual version requires integration over the (boundary) sphere, a topic deferred to the next chapter. See also Exercises 6 and 7 in that chapter.

Lemma 4.5 *We have the identity*

$$\int_B (v\triangle u - u\triangle v)\eta\, dx = \int_B u(\nabla v \cdot \nabla \eta) - v(\nabla u \cdot \nabla \eta)\, dx.$$

Here ∇u is the **gradient** of u, that is, $\nabla u = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_d}\right)$ and

$$\nabla v \cdot \nabla \eta = \sum_{j=1}^{d} \frac{\partial v}{\partial x_j}\frac{\partial \eta}{\partial x_j},$$

with $\nabla u \cdot \nabla \eta$ defined similarly.

In fact, by integrating by parts as in the proof of (14) we have

$$\int_B \frac{\partial u}{\partial x_j} v\eta\, dx = -\int_B u\frac{\partial v}{\partial x_j}\eta\, dx - \int_B uv\frac{\partial \eta}{\partial x_j}\, dx.$$

We then repeat this with u replaced by $\partial u/\partial x_j$, and sum in j to obtain

$$\int_B (\triangle u)v\eta\, dx = -\int_B (\nabla u \cdot \nabla v)\eta\, dx - \int_B (\nabla u \cdot \nabla \eta)v\, dx.$$

This yields the lemma if we subtract from this the symmetric formula with u and v interchanged.

We shall apply the lemma when u is a given harmonic function, while v is one of the three following "test" functions: first, $v(x) = 1$; second, $v(x) = |x|^2$; and third, $v(x) = |x|^{-d+2}$ if $d \geq 3$, while $v(x) = \log|x|$ if $d = 2$. The relevance of these choices arises because $\triangle v = 0$ in the first case, while $\triangle v$ is a non-zero constant in the second case; also v in the third case is a constant multiple of a "fundamental solution," and in particular $v(x)$ is harmonic for $x \neq 0$.

When $v(x) = 1$, we take $\eta = \eta_\epsilon^+$, where $\eta_\epsilon^+(x) = 1$ for $|x| \leq 1 - \epsilon$, $\eta_\epsilon^+(x) = 0$ for $|x| \geq 1$, and $|\nabla \eta_\epsilon^+(x)| \leq c/\epsilon$. We accomplish this by setting $\eta_\epsilon^+(x) = \chi\left(\frac{|x|-1+\epsilon}{\epsilon}\right)$ for $1 - \epsilon \leq |x| \leq 1$, where χ is a fixed C^2 function on $[0,1]$ that equals 1 in $[0,1/4]$ and equals 0 in $[3/4,1]$. A picture of η_ϵ^+ is given in Figure 3.

Since u is harmonic, we see that with $v = 1$, Lemma 4.5 implies

$$\int_B \nabla u \cdot \nabla \eta_\epsilon^+\, dx = 0. \tag{22}$$

Next we take $v(x) = |x|^2$; then clearly $\triangle v = 2d$, and with $\eta = \eta_\epsilon^+$ the lemma yields:

$$2d\int_B u\eta_\epsilon^+\, dx = \int_B |x|^2(\nabla u \cdot \nabla \eta_\epsilon^+)\, dx - 2\int_B u(x \cdot \nabla \eta_\epsilon^+)\, dx.$$

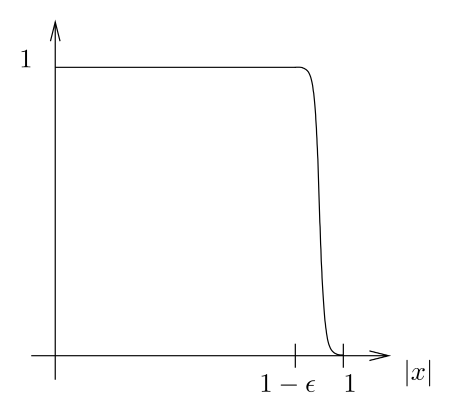


Figure 3. The function η_ϵ^+

However, since $\nabla\eta_\epsilon^+$ is supported in the spherical shell $S_\epsilon^+ = \{x : 1-\epsilon \leq |x| \leq 1\}$, we see that

$$\int_B |x|^2(\nabla u \cdot \nabla \eta_\epsilon^+)\, dx = \int_B (\nabla u \cdot \nabla \eta_\epsilon^+)\, dx + O(\epsilon),$$

and hence by (22) we get

$$d\int_B u\, dx = -\lim_{\epsilon\to 0}\int_B u(x\cdot\nabla\eta_\epsilon^+)\, dx. \tag{23}$$

We finally turn to $v(x) = |x|^{-d+2}$, when $d \geq 3$, and calculate $(\triangle v)(x)$ for $x \neq 0$ to see that it vanishes there. In fact, since $\partial|x|/\partial x_j = x_j/|x|$, we note that

$$\frac{\partial|x|^a}{\partial x_j} = ax_j|x|^{a-2} \quad \text{and} \quad \frac{\partial^2|x|^a}{\partial x_j^2} = a|x|^{a-2} + a(a-2)x_j^2|x|^{a-4}.$$

Upon adding in j, we obtain that $\triangle(|x|^a) = [da + a(a-2)]|x|^{a-2}$, and this is zero if $a = -d+2$ (or $a = 0$). A similar argument shows that $\triangle(\log|x|) = 0$ when $d = 2$ and $x \neq 0$.

We now apply the lemma with this v and $\eta = \eta_\epsilon$ defined as follows:

$$\begin{aligned} &\eta_\epsilon(x) = 1 - \chi(|x|/\epsilon) && \text{for } |x| \leq \epsilon, \\ &\eta_\epsilon(x) = 1 && \text{for } \epsilon \leq |x| \leq 1-\epsilon, \\ &\eta_\epsilon(x) = \eta_\epsilon^+(x) = \chi\left(\tfrac{|x|-1+\epsilon}{\epsilon}\right) && \text{for } 1-\epsilon \leq |x| \leq 1. \end{aligned}$$

The picture for η_ϵ is as follows (Figure 4):

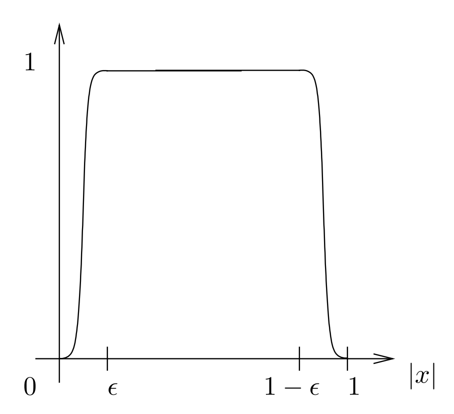


Figure 4. The function η_ϵ

We note that $|\nabla \eta_\epsilon|$ is $O(1/\epsilon)$ throughout. Now both u and v are harmonic in the support of η_ϵ, and in this case $\nabla \eta_\epsilon$ is supported only near the unit sphere (in the shell S_ϵ^+) or near the origin (in the ball $B_\epsilon = \{|x| < \epsilon\}$). Thus the right-hand side of the identity of the lemma gives two contributions, one over S_ϵ^+ and the other over B_ϵ. We consider the first contribution (when $d \geq 3$); it is

$$\int_{S_\epsilon^+} u\nabla(|x|^{-d+2}) \cdot \nabla \eta_\epsilon \, dx - \int_{S_\epsilon^+} |x|^{-d+2}(\nabla u \cdot \nabla \eta_\epsilon^+) \, dx.$$

Now the first integral is $(-d+2)\int_{S_\epsilon^+} u|x|^{-d}(x \cdot \nabla \eta_\epsilon^+) \, dx$, which by (23) tends to $c\int_B u \, dx$ as $\epsilon \to 0$, where c is the constant $(2-d)d$, since $|x|^{-d} - 1 = O(\epsilon)$ over S_ϵ^+. The second term tends to zero as $\epsilon \to 0$ because of (22) and the fact that the integrand there is supported in the shell S_ϵ^+. A similar argument for $d = 2$, with $v(x) = \log|x|$, yields the result with $c = 1$.

To consider the contribution near the origin, that is, over B_ϵ, we temporarily make the additional assumption that $u(0) = 0$. Then because of the differentiability assumption satisfied by a harmonic function, we have $u(x) = O(|x|)$ as $|x| \to 0$. Now over B_ϵ we have two terms, the first being $\int_{B_\epsilon} u\nabla(|x|^{-d+2})\nabla \eta_\epsilon \, dx$, which is majorized by

$$\int_{B_\epsilon} O(\epsilon)|x|^{-d+1}O(1/\epsilon) \, dx \leq O\left(\int_{|x|\leq\epsilon} |x|^{-d+1} \, dx\right) \leq O(\epsilon),$$

because of (8) in Section 2 of Chapter 2. This term tends to 0 with ϵ.

The second term is $\int_{B_\epsilon} |x|^{-d+2}(\nabla u \cdot \nabla \eta_\epsilon)\, dx$, which is majorized by

$$\frac{c_1}{\epsilon} \int_{|x|\leq\epsilon} |x|^{-d+2} = c_2\epsilon,$$

using the result just cited. We have used the fact that ∇u is bounded and $\nabla \eta_\epsilon$ is $O(1/\epsilon)$ throughout B. Letting $\epsilon \to 0$ we see that this term tends to zero also. A similar argument works when $d = 2$.

Thus we have proved that if u is harmonic in a neighborhood of the closure of the unit ball B, and $u(0) = 0$, then $\int_B u\, dx = 0$. We can drop the assumption $u(0) = 0$ by applying the conclusion we have just reached to $u(x) - u(0)$ in place of $u(x)$. Therefore we have achieved the mean-value property (21) for the unit ball.

Now suppose $B_r(x_0) = \{x : |x - x_0| < r\}$ is the ball of radius r centered at x_0, and consider $U(x) = u(x_0 + rx)$. If we suppose that u is harmonic in $B_r(x_0)$, then clearly U is harmonic in the unit ball (indeed, the property of being harmonic is unchanged under translations $x \to x + x_0$ and dilations $x \to rx$, as is easily verified). Thus if u were supported in Ω, and $B_r(x_0) \subset \Omega$, then by the result just proved $U(0) = \frac{1}{m(B)} \int_B U(x)\, dx$, which means that

$$\begin{aligned} u(x_0) &= \frac{1}{m(B)} \int_{|x|\leq 1} u(x_0 + rx)\, dx = \frac{1}{r^d m(B)} \int_{|x|\leq r} u(x_0 + x)\, dx \\ &= \frac{1}{m(B_r(x_0))} \int_{B_r} u(x)\, dx, \end{aligned}$$

by the relative invariance of Lebesgue measure under dilations and translations. This establishes (21) in general.

The converse property

To prove this, we first show that the mean-value property allows a useful extension of itself. For this purpose, we fix a function $\varphi(y)$ that is continuous in the closed unit ball $\{|y| \leq 1\}$ and is radial (that is, $\varphi(y) = \Phi(|y|)$ for an appropriate Φ), and extend φ to be zero when $|y| > 1$. Suppose in addition that $\int \varphi(y)\, dy = 1$. We then claim the following:

Lemma 4.6 *Whenever u satisfies the mean-value property* (21) *in Ω, and the closure of the ball $\{x : |x - x_0| < r\}$ lies in Ω, then*

(24)

$$u(x_0) = \int_{\mathbb{R}^d} u(x_0 - ry)\varphi(y)\, dy = \int_{\mathbb{R}^d} u(x_0 - y)\varphi_r(y)\, dy = (u * \varphi_r)(x_0),$$

where $\varphi_r(y) = r^{-d}\varphi(y/r)$.

That the second of the two identities holds is an immediate consequence of the change of variables $y \to y/r$; the rightmost equality is merely the definition of $u * \varphi_r$.

We can prove (24) as a consequence of a simple observation about integration. Let $\psi(y)$ be another function on the ball $\{|y| \leq 1\}$, which we assume is bounded. For each N, a large positive integer, denote by $B(j)$ the ball $\{|y| \leq j/N\}$. Recall that $\varphi(y) = \Phi(|y|)$. Then

$$\int \varphi(y)\psi(y)\,dy = \lim_{N\to\infty} \sum_{j=1}^{N} \Phi\left(\frac{j}{N}\right) \int_{B(j)-B(j-1)} \psi(y)\,dy. \tag{25}$$

To verify this, note that the left-hand side of (25) equals

$$\sum_{j=1}^{N} \int_{B(j)-B(j-1)} \varphi(y)\psi(y)\,dy.$$

However, $\sup_{1\leq j\leq N} \sup_{y\in B(j)-B(j-1)} |\varphi(y) - \Phi(j/N)| = \epsilon_N$, which tends to zero as $N \to \infty$, since φ is radial, continuous, and $\varphi(y) = \Phi(|y|)$. Thus the left-hand side of (25) differs from $\sum_{j=1}^{N} \Phi(j/N) \int_{B(j)-B(j-1)} \psi(y)\,dy$ by at most $\epsilon_N \int_{|y|\leq 1} |\psi(y)|\,dy$, proving (25).

We now use this in the case where $\psi(y) = u(x_0 - ry)$ and φ is as before. Then

$$\int u(x_0 - ry)\varphi(y)\,dy = \lim_{N\to\infty} \sum_{j=1}^{N} \Phi\left(\frac{j}{N}\right) \int_{B(j)-B(j-1)} u(x_0 - ry)\,dy.$$

However, it follows from the mean-value property assumed for u that

$$\int_{B(j)-B(j-1)} u(x_0 - ry)\,dy = u(x_0)[m(B(j)) - m(B(j-1))].$$

Therefore, the right-hand side above equals

$$u(x_0) \lim_{N\to\infty} \sum_{j=1}^{N} \Phi\left(\frac{j}{N}\right) \int_{B(j)-B(j-1)} dy,$$

and this is $u(x_0)$ if we use (25) again, this time with $\psi = 1$, and recall that $\int \varphi(y)\,dy = 1$. We have therefore proved the lemma.

We see from this that every continuous function which satisfies the mean-value property is its own regularization! To be precise, we have

$$u(x) = (u * \varphi_r)(x) \tag{26}$$

whenever $x \in \Omega$ and the distance from x to the boundary of Ω is larger than r. If we now require in addition that $\varphi \in C_0^\infty\{|y| < 1\}$, then by the discussion in Section 1 we conclude that u is smooth throughout Ω.

Let us now establish that such functions are harmonic. Indeed, by Taylor's theorem, for every $x_0 \in \Omega$

$$u(x_0 + x) - u(x_0) = \sum_{j=1}^{d} a_j x_j + \frac{1}{2} \sum_{j,k=1}^{d} a_{jk} x_j x_k + \epsilon(x), \tag{27}$$

where $\epsilon(x) = O(|x|^3)$ as $|x| \to 0$. We note next that $\int_{|x|\leq r} x_j\, dx = 0$ and $\int_{|x|\leq r} x_j x_k\, dx = 0$ for all j and k with $k \neq j$. This follows by carrying out the integrations first in the x_j variable and noting that the integral vanishes because x_j is an odd function. Also by an obvious symmetry $\int_{|x|\leq r} x_j^2\, dx = \int_{|x|\leq r} x_k^2\, dx$, and by the relative dilation-invariance (see Section 3, Chapter 1) these are equal to $r^2 \int_{|x|\leq r} (x_1/r)^2\, dx = r^{d+2} \int_{|x|\leq 1} x_1^2\, dx = cr^{d+2}$, with $c > 0$. We now integrate both sides of (27) over the ball $\{|x| \leq r\}$, divide by r^d, and use the mean-value property. The result is that

$$\frac{c}{2} r^2 \sum_{j=1}^{d} a_{jj} = \frac{cr^2}{2} (\triangle u)(x_0) = O\left(\frac{1}{r^d} \int_{|x|\leq r} |\epsilon(x)|\, dx\right) = O(r^3).$$

Letting $r \to 0$ then gives $\triangle u(x_0) = 0$. Since x_0 was an arbitrary point of Ω, the proof of Theorem 4.2 is concluded.

Theorem 4.3 and some corollaries

We come now to the proof of Theorem 4.3. Let us assume that u is weakly harmonic in Ω. For each $\epsilon > 0$ we define Ω_ϵ to be the set of points in Ω that are at a distance greater than ϵ from its boundary:

$$\Omega_\epsilon = \{x \in \Omega : \ d(x, \partial\Omega) > \epsilon\}.$$

Notice that Ω_ϵ is open, and that every point of Ω belongs to Ω_ϵ if ϵ is small enough. Then the regularization $u * \varphi_r = u_r$ considered in the previous theorem is defined in Ω_ϵ, for $r < \epsilon$, and as we have noted is a smooth function there. We next observe that it is weakly harmonic in

Ω_ϵ. In fact, for $\psi \in C_0^\infty(\Omega_\epsilon)$ we have

$$\begin{aligned}(u_r, \triangle\psi) &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} u(x-ry)\varphi(y)\,dy\right)(\triangle\psi)(x)\,dx \\ &= \int_{\mathbb{R}^d}\varphi(y)\left(\int_{\mathbb{R}^d} u(x-ry)(\triangle\psi)(x)\,dx\right)dy,\end{aligned}$$

by Fubini's theorem, and the inner integral vanishes for y, $|y| \leq 1$, because it equals $(u, \triangle\psi_r)$, with $\psi_r = \psi(x+ry)$. Thus we have

$$(u * \varphi_r, \triangle\psi) = 0,$$

and hence $u * \varphi_r$ is weakly harmonic. Next, since this regularization is automatically smooth it is then also harmonic. Moreover, we claim that

$$(u * \varphi_{r_1})(x) = (u * \varphi_{r_2})(x) \tag{28}$$

whenever $x \in \Omega_\epsilon$ and $r_1 + r_2 < \epsilon$. Indeed, $(u * \varphi_{r_1}) * \varphi_{r_2} = u * \varphi_{r_1}$ as we have shown in (26) above. However convolutions are commutative (see Remark (6) in Chapter 2); thus $(u * \varphi_{r_1}) * \varphi_{r_2} = (u * \varphi_{r_2}) * \varphi_{r_1} = u * \varphi_{r_2}$, and (28) is proved.

Now we can let r_1 tend to zero, while keeping r_2 fixed. We know by the properties of approximations to the identity that $u * \varphi_{r_1}(x) \to u(x)$ for almost every x in Ω_ϵ; hence $u(x)$ equals $u_{r_2}(x)$ for almost every $x \in \Omega_\epsilon$. Thus u can be corrected on Ω_ϵ (setting it equal to u_{r_2}), so that it becomes harmonic there. Now since ϵ can be taken arbitrarily small, the proof of the theorem is complete.

We state several further corollaries arising out of the above theorems.

Corollary 4.7 *Every harmonic function is indefinitely differentiable.*

Corollary 4.8 *Suppose $\{u_n\}$ is a sequence of harmonic functions in Ω that converges to a function u uniformly on compact subsets of Ω as $n \to \infty$. Then u is also harmonic.*

The first of these corollaries was already proved as a consequence of (26). For the second, we use the fact that each u_n satisfies the mean-value property

$$u_n(x_0) = \frac{1}{m(B)}\int_B u_n(x)\,dx$$

whenever B is a ball with center at x_0, and $\overline{B} \subset \Omega$. Thus by the uniform convergence it follows that u also satisfies this property, and hence u is harmonic.

We should point out that these properties of harmonic functions on $\mathbb{R}^d$ are reminiscent of similar properties of holomorphic functions. But this should not be surprising, given the close connection between these two classes of functions in the special case $d = 2$.

4.2 The boundary value problem and Dirichlet's principle

The d-dimensional Dirichlet boundary value problem we are concerned with may be stated as follows. Let Ω be an open bounded set in $\mathbb{R}^d$. Given a continuous function f defined on the boundary $\partial\Omega$, we wish to find a function u that is continuous in $\overline{\Omega}$, harmonic in Ω, and such that $u = f$ on $\partial\Omega$.

An important preliminary observation is that the solution to the problem, if it exists, is unique. Indeed, if u_1 and u_2 are two solutions then $u_1 - u_2$ is harmonic in Ω and vanishes on the boundary. Thus by the maximum principle (Corollary 4.4) we have $u_1 - u_2 = 0$, and hence $u_1 = u_2$.

Turning to the existence of a solution, we shall now pursue the approach of Dirichlet's principle outlined earlier.

We consider the class of functions $C^1(\overline{\Omega})$, and equip this space with the inner product

$$\langle u, v\rangle = \int_\Omega (\nabla u \cdot \overline{\nabla v})\, dx,$$

where of course

$$\nabla u \cdot \overline{\nabla v} = \sum_{j=1}^{d} \frac{\partial u}{\partial x_j}\overline{\frac{\partial v}{\partial x_j}}.$$

With this inner product, we have a corresponding norm given by $\|u\|^2 = \langle u, u\rangle$. We note that $\|u\| = 0$ is the same as $\nabla u = 0$ throughout Ω, which means that u is constant on each connected component of Ω. Thus we are led to consider equivalence classes in $C^1(\overline{\Omega})$ of elements modulo functions that are constant on components of Ω. These then form a pre-Hilbert space with inner product and norm given as above. We call this pre-Hilbert space $\mathcal{H}_0$.

In studying the completion $\mathcal{H}$ of $\mathcal{H}_0$ and its applications to the boundary value problem, the following lemma is needed.

Lemma 4.9 *Let Ω be an open bounded set in $\mathbb{R}^d$. Suppose v belongs to $C^1(\overline{\Omega})$ and v vanishes on $\partial\Omega$. Then*

$$\int_\Omega |v(x)|^2\, dx \leq c_\Omega \int_\Omega |\nabla v(x)|^2\, dx. \tag{29}$$

Proof. This conclusion could in fact be deduced from the considerations given in Lemma 3.3. We prefer to prove this easy version separately to highlight a simple idea that we shall also use later. It should be noted that the argument yields the estimate $c_\Omega \leq d(\Omega)^2$, where $d(\Omega)$ is the diameter of Ω.

We proceed on the basis of the following observation. Suppose f is a function in $C^1(\overline{I})$, where $I = (a, b)$ is an interval in $\mathbb{R}$. Assume that f vanishes at *one* of the end-points of I. Then

$$\int_I |f(t)|^2 \, dt \leq |I|^2 \int_I |f'(t)|^2 \, dt, \tag{30}$$

where $|I|$ denotes the length of I.

Indeed, suppose $f(a) = 0$. Then $f(s) = \int_a^s f'(t)\, dt$, and by the Cauchy-Schwarz inequality

$$|f(s)|^2 \leq |I| \int_a^s |f'(t)|^2 \, dt \leq |I| \int_I |f'(t)|^2 \, dt.$$

Integrating this in s over I then yields (30).

To prove (29), write $x = (x_1, x')$ with $x_1 \in \mathbb{R}$ and $x' \in \mathbb{R}^{d-1}$ and apply (30) to f defined by $f(x_1) = v(x_1, x')$, with x' fixed. Let $J(x')$ be the open set in $\mathbb{R}$ that is the corresponding slice of Ω given by $\{x_1 \in \mathbb{R} : (x_1, x') \in \Omega\}$. The set $J(x')$ can be written as a disjoint union of open intervals I_j. (Note that in fact $f(x_1)$ vanishes at both end-points of each I_j.) For each j, on applying (30) we obtain

$$\int_{I_j} |v(x_1, x')|^2 \, dx_1 \leq |I_j|^2 \int_{I_j} |\nabla v(x_1, x')|^2 \, dx_1.$$

Now since $|I_j| \leq d(\Omega)$, summing over the disjoint intervals I_j gives

$$\int_{J(x')} |v(x_1, x')|^2 \, dx_1 \leq d(\Omega)^2 \int_{J(x')} |\nabla v(x_1, x')|^2 \, dx_1,$$

and an integration over $x' \in \mathbb{R}^d$ then leads to (29).

Now let S_0 denote the linear subspace of $C^1(\overline{\Omega})$ consisting of functions that vanish on the boundary of Ω. We note that distinct elements of S_0 remain distinct under the equivalence relation defining $\mathcal{H}_0$ (since constants on each component that vanish on the boundary are zero), and so S_0 may be identified with a subspace of $\mathcal{H}_0$. Denote by S the closure in $\mathcal{H}$ of this subspace, and let P_S be the orthogonal projection of $\mathcal{H}$ onto S.

With these preliminaries out of the way, we first try to solve the boundary value problem with f given on $\partial\Omega$ under the additional assumption that f is the restriction to $\partial\Omega$ of a function F in $C^1(\overline{\Omega})$. (How this additional hypothesis can be removed will be explained below.) Following the prescription of Dirichlet's principle, we seek a sequence $\{u_n\}$ with $u_n \in C^1(\overline{\Omega})$ and $u_n|_{\partial\Omega} = F|_{\partial\Omega}$, such that the Dirichlet integrals $\|u_n\|^2$ converge to a minimum value. This means that $u_n = F - v_n$, with $v_n \in S_0$, and that $\lim_{n\to\infty} \|u_n\|$ minimizes the distance from F to S_0. Since $S = \overline{S_0}$, this sequence also minimizes the distance from F to S in $\mathcal{H}$.

Now what do the elementary facts about orthogonal projections teach us? According to the proof of Lemma 4.1 in the previous chapter, we conclude that the sequence $\{v_n\}$, and hence also the sequence $\{u_n\}$, both converge in the norm of $\mathcal{H}$, the former having a limit $P_S(F)$. Now applying Lemma 4.9 to $v_n - v_m$ we deduce that $\{v_n\}$ and $\{u_n\}$ are also Cauchy in the $L^2(\Omega)$-norm, and thus converge also in the L^2-norm. Let $u = \lim_{n\to\infty} u_n$. Then

$$u = F - P_S(F). \tag{31}$$

We see that u is weakly harmonic. Indeed, whenever $\psi \in C_0^\infty(\Omega)$, then $\psi \in S$, and hence by (31) $\langle u, \psi\rangle = 0$. Therefore $\langle u_n, \psi\rangle \to 0$, but by integration by parts, as we have seen,

$$\langle u_n, \psi\rangle = \int_\Omega (\nabla u_n \cdot \overline{\nabla\psi})\, dx = -\int_\Omega u_n \overline{\triangle\psi}\, dx = -(u_n, \triangle\psi).$$

As a result, $(u, \triangle\psi) = 0$, and so u is weakly harmonic and thus can be corrected on a set of measure zero to become harmonic.

This is the purported solution to our problem. However, two issues still remain to be resolved.

The first is that while u is the limit of a sequence $\{u_n\}$ of continuous functions in $\overline{\Omega}$ and $u_n|_{\partial\Omega} = f$, for each n, it is not clear that u itself is continuous in $\overline{\Omega}$ and $u|_{\partial\Omega} = f$.

The second issue is that we restricted our argument above to those f defined on the boundary of Ω that arise as restrictions of functions in $C^1(\overline{\Omega})$.

The second obstacle is the easier of the two to overcome, and this can be done by the use of the following lemma, applied to the set $\Gamma = \partial\Omega$.

Lemma 4.10 *Suppose Γ is a compact set in $\mathbb{R}^d$, and f is a continuous function on Γ. Then there exists a sequence $\{F_n\}$ of smooth functions on $\mathbb{R}^d$ so that $F_n \to f$ uniformly on Γ.*

In fact, supposing we can deal with the first issue raised, then with the lemma we proceed as follows. We find the functions U_n that are harmonic in Ω, continuous on $\overline{\Omega}$, and such that $U_n|_{\partial\Omega} = F_n|_{\partial\Omega}$. Now since the $\{F_n\}$ converges uniformly (to f) on $\partial\Omega$, it follows by the maximum principle that the sequence $\{U_n\}$ converges uniformly to a function u that is continuous on $\overline{\Omega}$, has the property that $u|_{\partial\Omega} = f$, and which is moreover harmonic (by Corollary 4.8 above). This achieves our goal.

The proof of Lemma 4.10 is based on the following extension principle.

Lemma 4.11 *Let f be a continuous function on a compact subset Γ of $\mathbb{R}^d$. Then there exists a function G on $\mathbb{R}^d$ that is continuous, and so that $G|_{\partial\Gamma} = f$.*

Proof. We begin with the observation that if K_0 and K_1 are two disjoint compact sets, there exists a continuous function $0 \leq g(x) \leq 1$ on $\mathbb{R}^d$ which takes the value 0 on K_0 and 1 on K_1. Indeed, if $d(x, \Omega)$ denotes the distance from x to Ω, we see that

$$g(x) = \frac{d(x, K_0)}{d(x, K_0) + d(x, K_1)}$$

has the required properties.

Now, we may assume without loss of generality that f is non-negative and bounded by 1 on Γ. Let

$$K_0 = \{x \in \Gamma : 2/3 \leq f(x) \leq 1\} \quad \text{and} \quad K_1 = \{x \in \Gamma : 0 \leq f(x) \leq 1/3\},$$

so that K_0 and K_1 are disjoint. Clearly, the observation before the lemma guarantees that there exists a function $0 \leq G_1(x) \leq 1/3$ on $\mathbb{R}^d$ which takes the value $1/3$ on K_0 and 0 on K_1. Then we see that

$$0 \leq f(x) - G_1(x) \leq \frac{2}{3} \quad \text{for all } x \in \Gamma.$$

We now repeat the argument with f replaced by $f - G_1$. In the first step, we have gone from $0 \leq f \leq 1$ to $0 \leq f - G_1 \leq 2/3$. Consequently, we may find a continuous function G_2 on $\mathbb{R}^d$ so that

$$0 \leq f(x) - G_1(x) - G_2(x) \leq \left(\frac{2}{3}\right)^2 \quad \text{on } \Gamma,$$

and $0 \leq G_2 \leq \frac{1}{3}\frac{2}{3}$. Repeating this process, we find continuous functions G_n on $\mathbb{R}^d$ such that

$$0 \leq f(x) - G_1(x) - \cdots - G_N(x) \leq \left(\frac{2}{3}\right)^N \quad \text{on } \Gamma,$$

and $0 \leq G_N \leq \frac{1}{3}\left(\frac{2}{3}\right)^{N-1}$ on $\mathbb{R}^d$. If we define

$$G = \sum_{n=1}^{\infty} G_n,$$

then G is continuous and equals f on Γ.

To complete the proof of Lemma 4.10, we argue as follows. We regularize the function G obtained in Lemma 4.11 by defining

$$F_\epsilon(x) = \epsilon^{-d} \int_{\mathbb{R}^d} G(x-y)\varphi(y/\epsilon)\,dy = \int_{\mathbb{R}^d} G(y)\varphi_\epsilon(x-y)\,dy,$$

with $\varphi_\epsilon(y) = \epsilon^{-d}\varphi(y/\epsilon)$, where φ is a non-negative C_0^∞ function supported in the unit ball with $\int \varphi(y)\,dy = 1$. Then each F_ϵ is a C^∞ function. However,

$$F_\epsilon(x) - G(x) = \int (G(y) - G(x))\varphi_\epsilon(x-y)\,dy.$$

Since the integration above is restricted to $|x-y| \leq \epsilon$, then if $x \in \Gamma$, we see that

$$\begin{aligned} |F_\epsilon(x) - G(x)| &\leq \sup_{|x-y|\leq\epsilon} |G(x) - G(y)| \int \varphi_\epsilon(x-y)\,dy \\ &\leq \sup_{|x-y|\leq\epsilon} |G(x) - G(y)|. \end{aligned}$$

The last quantity tends to zero with ϵ by the uniform continuity of G near Γ, and if we choose $\epsilon = 1/n$ we obtain our desired sequence.

The two-dimensional theorem

We now take up the problem of whether the proposed solution u takes on the desired boundary values. Here we limit our discussion to the case of two dimensions for the reason that in the higher dimensional situation the problems that arise involve a number of questions that would take us beyond the scope of this book. In contrast, in two dimensions, while the proof of the result below is a little tricky, it is within the reach of the Hilbert space methods we have been illustrating.

The Dirichlet problem can be solved (in two dimensions as well as in higher dimensions) only if certain restrictions are made concerning the nature of the domain Ω. The regularity we shall assume, while not

optimal,[15] is broad enough to encompass many applications, and yet has a simple geometric form. It can be described as follows. We fix an initial triangle T_0 in $\mathbb{R}^2$. To be precise, we assume that T_0 is an isosceles triangle whose two equal sides have length ℓ, and make an angle α at their common vertex. The exact values of ℓ and α are unimportant; they may both be taken as small as one wishes, but must be kept fixed throughout our discussion. With the shape of T_0 thus determined, we say that T is a **special triangle** if it is congruent to T_0, that is, T arises from T_0 by a translation and rotation. The **vertex** of T is defined to be the intersection of its two equal sides.

The regularity property of Ω we assume, the **outside-triangle condition**, is as follows: with ℓ and α fixed, for each x in the boundary of Ω, there is a special triangle with vertex x whose interior lies outside Ω. (See Figure 5.)

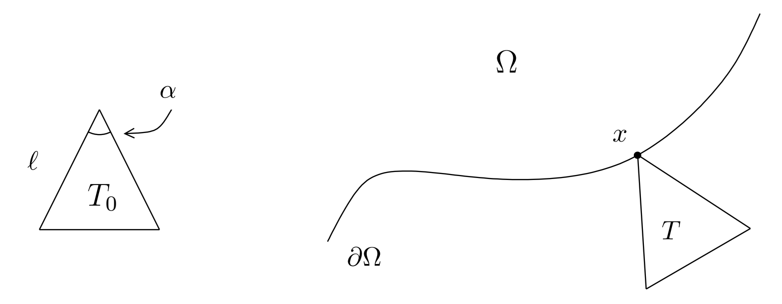


Figure 5. The triangle T_0 and the special triangle T

Theorem 4.12 *Let Ω be an open bounded set in $\mathbb{R}^2$ that satisfies the outside-triangle condition. If f is a continuous function on $\partial\Omega$, then the boundary value problem $\triangle u = 0$ with u continuous in $\overline{\Omega}$ and $u|_{\partial\Omega} = f$ is always uniquely solvable.*

Some comments are in order.

(1) If Ω is bounded by a polygonal curve, it satisfies the conditions of the theorem.

(2) More generally, if Ω is appropriately bounded by finitely many Lipschitz curves, or in particular C^1 curves, the conditions are also satisfied.

(3) There are simple examples where the problem is not solvable: for instance, if Ω is the punctured disc. This example of course does not satisfy the outside-triangle condition.

---

[15] The optimal conditions involve the notion of capacity of sets.

(4) The conditions on Ω in this theorem are not optimal: one can construct examples of Ω when the problem is solvable for which the above regularity fails.

For more details on the above, see Exercise 19 and Problem 4.

We turn to the proof of the theorem. It is based on the following proposition, which may be viewed as a refined version of Lemma 4.9 above.

Proposition 4.13 *For any bounded open set Ω in $\mathbb{R}^2$ that satisfies the outside-triangle condition there are two constants $c_1 < 1$ and $c_2 > 1$ such that the following holds. Suppose z is a point in Ω whose distance from $\partial\Omega$ is δ. Then whenever v belongs to $C^1(\overline{\Omega})$ and $v|_{\partial\Omega} = 0$, we have*

$$\int_{B_{c_1\delta}(z)} |v(x)|^2\, dx \leq C\delta^2 \int_{B_{c_2\delta}(z)\cap\Omega} |\nabla v(x)|^2\, dx. \tag{32}$$

The bound C can be chosen to depend only on the diameter of Ω and the parameters ℓ and α which determine the triangles T.

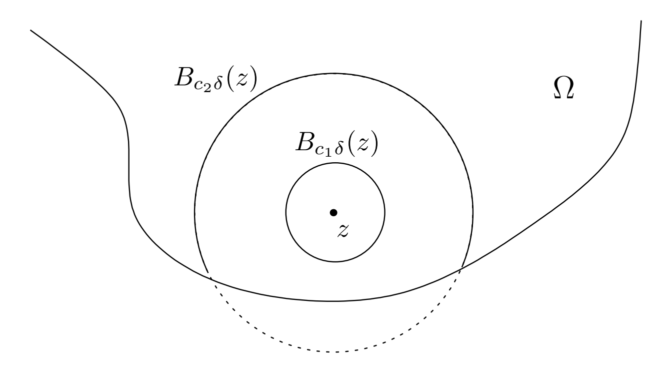


Figure 6. The situation in Proposition 4.13

Let us see how the proposition proves the theorem. We have already shown that it suffices to assume that f is the restriction to $\partial\Omega$ of an F that belongs to $C^1(\overline{\Omega})$. We recall we had the minimizing sequence $u_n = F - v_n$, with $v_n \in C^1(\overline{\Omega})$ and $v_n|_{\partial\Omega} = 0$. Moreover, this sequence converges in the norm of $\mathcal{H}$ and $L^2(\Omega)$ to a limit v, such that $u = F - v$ is harmonic in Ω. Then since (32) holds for each v_n, it also holds for $v = F - u$; that is,

$$\int_{B_{c_1\delta}(z)} |(F-u)(x)|^2\, dx \leq C\delta^2 \int_{B_{c_2\delta}(z)\cap\Omega} |\nabla(F-u)(x)|^2\, dx. \tag{33}$$

To prove the theorem it suffices, in view of the continuity of u in Ω, to show that if y is any fixed point in $\partial\Omega$, and z is a variable point in Ω, then $u(z) \to f(y)$ as $z \to y$. Let $\delta = \delta(z)$ denote the distance of z from the boundary. Then $\delta(z) \leq |z - y|$ and therefore $\delta(z) \to 0$ as $z \to y$.

We now consider the averages of F and u taken over the discs centered at z of radius $c_1\delta(z)$ (recall that $c_1 < 1$). We denote these averages by $\mathrm{Av}(F)(z)$ and $\mathrm{Av}(u)(z)$, respectively. Then by the Cauchy-Schwarz inequality, we have

$$|\mathrm{Av}(F)(z) - \mathrm{Av}(u)(z)|^2 \leq \frac{1}{\pi(c_1\delta)^2} \int_{B_{c_1\delta}(z)\cap\Omega} |F - u|^2 \, dx,$$

which by (33) is then majorized by

$$C' \int_{B_{c_2\delta}(z)\cap\Omega} |\nabla(F - u)|^2 \, dx.$$

The absolute continuity of the integral guarantees that the last integral tends to zero with δ, since $m(B_{c_2\delta}) \to 0$. However, by the mean-value property, $\mathrm{Av}(u)(z) = u(z)$, while by the continuity of F in $\overline{\Omega}$,

$$\mathrm{Av}(F)(z) = \frac{1}{m(B_{c_1\delta}(z))} \int_{B_{c_1\delta}(z)} F(x) \, dx \to f(y),$$

because $F|_{\partial\Omega} = f$ and $z \to y$. Altogether this gives $u(z) \to f(y)$, and the theorem is proved, once the proposition is established.

To prove the proposition, we construct for each $z \in \Omega$ whose distance from $\partial\Omega$ is δ, and for δ sufficiently small, a rectangle R with the following properties:

(1) R has side lengths $2c_1\delta$ and $M\delta$ (with $c_1 \leq 1/2$, $M \leq 4$).

(2) $B_{c_1\delta}(z) \subset R$.

(3) Each segment in R, that is parallel to and of length equal to the length of the long side, intersects the boundary of Ω.

To obtain R we let y be a point in $\partial\Omega$ so that $\delta = |z - y|$, and we apply the outside-triangle condition at y. As a result, the line joining z with y and *one* of the sides of the special triangle whose vertex is at y must make an angle $\beta < \pi$. (In fact $\beta \leq \pi - \alpha/2$, as is easily seen.) Now after a suitable rotation and translation we may assume that $y = 0$ and that the angle going from the x_2-axis to the line joining z to 0 is equal to the

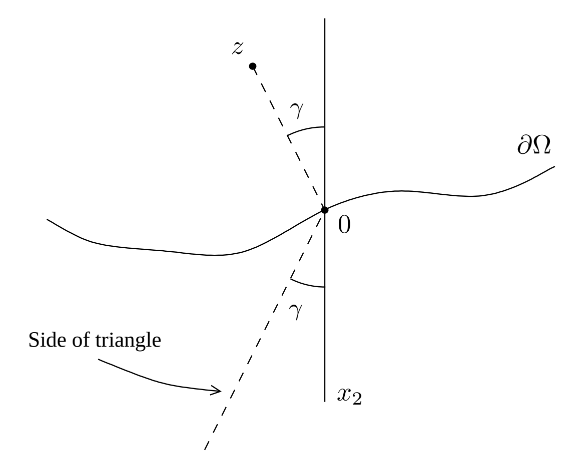


Figure 7. Placement of the rectangle R

angle of the side of the triangle to the x_2-axis. This angle can be taken to be γ, with $\gamma > \alpha/4$. (See Figure 7.)

There is an alternate possibility that occurs with this figure reflected through the x_2-axis.

With this picture in mind we construct the rectangle R as indicated in Figure 8.

It has its long side parallel to the x_2-axis, contains the disc $B_{c_1\delta}(z)$, and every segment R parallel to the x_2-axis intersects the (extension) of the side of the triangle.

Note that the coordinates of z are $(-\delta \sin\gamma, \delta\cos\gamma)$. We choose $c_1 < \sin\gamma$, then $B_{c_1\delta}(z)$ lies in the same (left) half-plane as z.

We next focus our attention on two points: P_1, which lies on the x_1-axis at the intersection of this axis with the far side of the rectangle; and P_2, which is at the corner of that side of the rectangle, that is, at the intersection of the (continuation) of the side of the outside triangle and the further side of the rectangle. The coordinates of P_1 are $(-a, 0)$, where $a = \delta c_1 + \delta \sin\gamma$. The coordinates of P_2 are $(-a, -a\frac{\cos\gamma}{\sin\gamma})$. Note that the distance of P_2 from the origin is $a/\sin\gamma$, which is $\delta + c_1\delta/\sin\gamma \leq 2\delta$, since $c_1 < \sin\gamma$.

Now we observe that the length of the larger side of the rectangle is the sum of the part that lies above the x_1-axis and the part that lies below. The upper part has length the sum of the radius of the disc plus the height of z, and this is $c_1\delta + \delta\cos\gamma \leq 2\delta$. The lower part has length equal to $a/\tan\gamma$, which is $\delta\cos\gamma + \delta c_1 \frac{\cos\gamma}{\sin\gamma} \leq 2\delta$, since $c_1 < \sin\gamma$. Thus

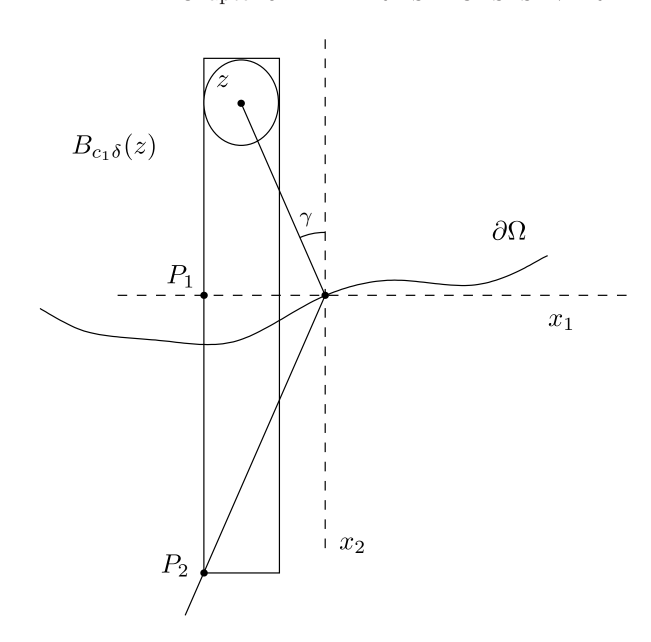


Figure 8. The disc $B_{c_1\delta}(z)$ and the rectangle R containing it

we find that the length of the side is $\leq 4\delta$.

Now it is clear from the construction that each vertical segment in R starting from the disc $B_{c_1\delta}(z)$ when continued downward and parallel to the x_2-axis intersects the line joining 0 to P_2, (which is a continuation of the side of the triangle). Moreover, if the length ℓ of this side of the triangle exceeds the distance of P_2 from the origin, then the segment intersects the triangle. When this intersection occurs the segment starting from $B_{c_2\delta}(z)$ must also intersect the boundary of Ω, since the triangle lies outside Ω. Therefore if $\ell \geq 2\delta$ the desired intersection occurs, and each of the conclusions (1), (2), and (3) are verified. (We shall lift the restriction $\delta \leq \ell/2$ momentarily.)

Now we integrate over each line segment parallel to the x_2-axis in R, including its portion in $B_{c_1\delta}(z)$, which is continued downward until it meets $\partial\Omega$. Call such a segment $I(x_1)$. Then, using (30) we see that

$$\int_{I(x_1)} |v(x_1, x_2)|^2 \, dx_2 \leq M^2\delta^2 \int_{I(x_1)} \left| \frac{\partial v}{\partial x_2}(x_1, x_2) \right|^2 dx_2,$$

and an integration in x_1 gives

$$\int_{R\cap\Omega} |v(x)|^2 \, dx \leq M\delta^2 \int_{R\cap\Omega} |\nabla v(x)|^2 \, dx.$$

However, we note that $B_{c_1\delta}(z) \subset R$, and $B_{c_2\delta}(z) \supset R$ when $c_2 \geq 2$. Thus the desired inequality (32) is established, still under the assumption that δ is small, that is, $\delta \leq \ell/2$. When $\delta > \ell/2$ it suffices merely to use the crude estimate (29) and the proposition is then proved. The proof of the theorem is therefore complete.

5 Exercises

1. Suppose $f \in L^2(\mathbb{R}^d)$ and $k \in L^1(\mathbb{R}^d)$.

(a) Show that $(f * k)(x) = \int f(x-y)k(y)\,dy$ converges for a.e. x.

(b) Prove that $\|f * k\|_{L^2(\mathbb{R}^d)} \leq \|f\|_{L^2(\mathbb{R}^d)}\|k\|_{L^1(\mathbb{R}^d)}$.

(c) Establish $\widehat{(f * k)}(\xi) = \hat{k}(\xi)\hat{f}(\xi)$ for a.e. ξ.

(d) The operator $Tf = f * k$ is a Fourier multiplier operator with multiplier $m(\xi) = \hat{k}(\xi)$.

[Hint: See Exercise 21 in Chapter 2.]

2. Consider the **Mellin transform** defined initially for continuous functions f of compact support in $\mathbb{R}^+ = \{t \in \mathbb{R} : \ t > 0\}$ and $x \in \mathbb{R}$ by

$$\mathcal{M}f(x) = \int_0^\infty f(t)t^{ix-1}dt.$$

Prove that $(2\pi)^{-1/2}\mathcal{M}$ extends to a unitary operator from $L^2(\mathbb{R}^+, dt/t)$ to $L^2(\mathbb{R})$. The Mellin transform serves on $\mathbb{R}^+$, with its multiplicative structure, the same purpose as the Fourier transform on $\mathbb{R}$, with its additive structure.

3. Let $F(z)$ be a bounded holomorphic function in the half-plane. Show in two ways that $\lim_{y\to 0} F(x+iy)$ exists for a.e. x.

(a) By using the fact that $F(z)/(z+i)$ is in $H^2(\mathbb{R}^2_+)$.

(b) By noting that $G(z) = F\left(i\frac{1-z}{1+z}\right)$ is a bounded holomorphic function in the unit disc, and using Exercise 17 in the previous chapter.

4. Consider $F(z) = e^{i/z}/(z+i)$ in the upper half-plane. Note that $F(x+iy) \in L^2(\mathbb{R})$, for each $y > 0$ and $y = 0$. Observe also that $F(z) \to 0$ as $|z| \to 0$. However, $F \notin H^2(\mathbb{R}^2_+)$. Why?

5. For $a < b$, let $S_{a,b}$ denote the strip $\{z = x+iy, \quad a < y < b\}$. Define $H^2(S_{a,b})$ to consist of the holomorphic functions F in $S_{a,b}$ so that

$$\|F\|^2_{H^2(S_{a,b})} = \sup_{a<y<b} \int_{\mathbb{R}^2} |F(x+iy)|^2\,dx < \infty.$$

Define $H^2(S_{a,\infty})$ and $H^2(S_{-\infty,b})$ to be the obvious variants of the Hardy spaces for the half-planes $\{z = x + iy,\ y > a\}$ and $\{z = x + iy,\ y < b\}$, respectively.

(a) Show that $F \in H^2(S_{a,b})$ if and only if F can be written as

$$F(z) = \int_{\mathbb{R}} f(\xi)e^{-2\pi i z\xi}\, d\xi,$$

with $\int_{\mathbb{R}} |f(\xi)|^2(e^{4\pi a\xi} + e^{4\pi b\xi})\, d\xi < \infty$.

(b) Prove that every $F \in H^2(S_{a,b})$ can be decomposed as $F = G_1 + G_2$, where $G_{\in} H^2(S_{a,\infty})$ and $G_2 \in H^2(S_{-\infty,b})$.

(c) Show that $\lim_{a<y<b, y\to a} F(x + iy) = F_a(x)$ exists in the L^2-norm and also almost everywhere, with a similar result for $\lim_{a<y<b, y\to b} F(x + iy)$.

6. Suppose Ω is an open set in $\mathbb{C} = \mathbb{R}^2$, and let $\mathcal{H}$ be the subspace of $L^2(\Omega)$ consisting of holomorphic functions on Ω. Show that $\mathcal{H}$ is a closed subspace of $L^2(\Omega)$, and hence is a Hilbert space with inner product

$$(f, g) = \int_{\Omega} f(z)\overline{g}(z)\, dx\, dy, \qquad \text{where } z = x + iy.$$

[Hint: Prove that for $f \in \mathcal{H}$, we have $|f(z)| \leq \frac{c}{d(z,\Omega^c)}\|f\|$ for $z \in \Omega$, where $c = \pi^{-1/2}$, using the mean-value property (9). Thus if $\{f_n\}$ is a Cauchy sequence in $\mathcal{H}$, it converges uniformly on compact subsets of Ω.]

7. Following up on the previous exercise, prove:

(a) If $\{\varphi_n\}_{n=0}^{\infty}$ is an orthonormal basis of $\mathcal{H}$, then

$$\sum_{n=0}^{\infty} |\varphi_n(z)|^2 \leq \frac{c^2}{d(z, \Omega^c)} \qquad \text{for } z \in \Omega.$$

(b) The sum

$$B(z, w) = \sum_{n=0}^{\infty} \varphi_n(z)\overline{\varphi_n}(w)$$

converges absolutely for $(z, w) \in \Omega \times \Omega$, and is independent of the choice of the orthonormal basis $\{\varphi_n\}$ of $\mathcal{H}$.

(c) To prove (b) it is useful to characterize the function $B(z, w)$, called the **Bergman kernel**, by the following property. Let T be the linear transformation on $L^2(\Omega)$ defined by

$$Tf(z) = \int_{\Omega} B(z, w)f(w)\, du\, dv, \qquad w = u + iv.$$

Then T is the orthogonal projection of $L^2(\Omega)$ to $\mathcal{H}$.

(d) Suppose that Ω is the unit disc. Then $f \in \mathcal{H}$ exactly when $f(z) = \sum_{n=0}^{\infty} a_n z^n$, with

$$\sum_{n=0}^{\infty} |a_n|^2 (n+1)^{-1} < \infty.$$

Also, the sequence $\{\frac{z^n(n+1)}{\pi^{1/2}}\}_{n=0}^{\infty}$ is an orthonormal basis of $\mathcal{H}$. Moreover, in this case

$$B(z,w) = \frac{1}{\pi(1 - z\overline{w})^2}.$$

8. Continuing with Exercise 6, suppose Ω is the upper half-plane $\mathbb{R}^2_+$. Then every $f \in \mathcal{H}$ has a representation

$$f(z) = \sqrt{4\pi} \int_0^{\infty} \hat{f}_0(\xi) e^{2\pi i \xi z}\, d\xi, \qquad z \in \mathbb{R}^2_+, \tag{34}$$

where $\int_0^\infty |\hat{f}_0(\xi)|^2 \frac{d\xi}{\xi} < \infty$. Moreover, the mapping $\hat{f}_0 \to f$ given by (34) is a unitary mapping from $L^2((0,\infty), \frac{d\xi}{\xi})$ to $\mathcal{H}$.

9. Let H be the Hilbert transform. Verify that

(a) $H^* = -H$, $H^2 = -I$, and H is unitary.

(b) If τ_h denotes the translation operator, $\tau_h(f)(x) = f(x-h)$, then H commutes with τ_h, $\tau_h H = H\tau_h$.

(c) If δ_a denotes the dilation operator, $\delta_a(f)(x) = f(ax)$ with $a > 0$, then H commutes with δ_a, $\delta_a H = H\delta_a$.

A converse is given in Problem 5 below.

10. Let $f \in L^2(\mathbb{R})$ and let $u(x,y)$ be the Poisson integral of f, that is $u = (f * \mathcal{P}_y)(x)$, as given in (10) above. Let $v(x,y) = (Hf * \mathcal{P}_y)(x)$, the Poisson integral of the Hilbert transform of f. Prove that:

(a) $F(x+iy) = u(x,y) + iv(x,y)$ is analytic in the half-plane $\mathbb{R}^2_+$, so that u and v are conjugate harmonic functions. We also have $f = \lim_{y\to 0} u(x,y)$ and $Hf = \lim_{y \to 0} v(x,y)$.

(b) $F(z) = \frac{1}{\pi i}\int_{\mathbb{R}} f(t) \frac{dt}{t-z}$.

(c) $v(x,y) = f * \mathcal{Q}_y$, where $\mathcal{Q}_y(x) = \frac{1}{\pi}\frac{x}{x^2+y^2}$ is the **conjugate Poisson kernel**.

[Hint: Note that $\frac{i}{\pi z} = \mathcal{P}_y(x) + i\mathcal{Q}_y(x)$, $z = x+iy$.]

11. Show that

$$\left\{ \frac{1}{\pi^{1/2}(i+z)} \left(\frac{i-z}{i+z}\right)^n \right\}_{n=0}^{\infty}$$

is an orthonormal basis of $H^2(\mathbb{R}^2_+)$.

Note that $\left\{\frac{1}{\pi^{1/2}(i+x)}\left(\frac{i-x}{i+x}\right)^n\right\}_{n=0}^{\infty}$ is an orthonormal basis of $L^2(\mathbb{R})$; see Exercise 9 in the previous chapter.

[Hint: It suffices to show that if $F \in H^2(\mathbb{R}^2_+)$ and

$$\int_{-\infty}^{\infty} F(x)\frac{(x+i)^n}{(x-i)^{n+1}}\,dx = 0 \qquad \text{for } n = 0,\, 1,\, 2,\, \ldots,$$

then $F = 0$. Use the Cauchy integral formula to prove that

$$\left(\frac{d}{dz}\right)^n (F(z)(z+i)^n)|_{z=i} = 0,$$

and thus $F^{(n)}(i) = 0$ for $n = 0, 1, 2, \ldots$.]

12. We consider whether the inequality

$$\|u\|_{L^2(\Omega)} \leq c\|L(u)\|_{L^2(\Omega)}$$

can hold for open sets Ω that are unbounded.

(a) Assume $d \geq 2$. Show that for each constant coefficient partial differential operator L, there are unbounded connected open sets Ω for which the above holds for all $u \in C_0^\infty(\Omega)$.

(b) Show that $\|u\|_{L^2(\mathbb{R}^d)} \leq c\|L(u)\|_{L^2(\mathbb{R}^d)}$ for all $u \in C_0^\infty(\mathbb{R}^d)$ if and only if $|P(\xi)| \geq c > 0$ all ξ, where P is the characteristic polynomial of L.

[Hint: For (a) consider first $L = (\partial/\partial x_1)^n$ and a strip $\{x: \ -1 < x_1 < 1\}$.]

13. Suppose L is a linear partial differential operator with constant coefficients. Show that when $d \geq 2$, the linear space of solutions u of $L(u) = 0$ with $u \in C^\infty(\mathbb{R}^d)$ is not finite-dimensional.

[Hint: Consider the zeroes ζ of $P(\zeta)$, $\zeta \in \mathbb{C}^d$, where P is the characteristic polynomial of L.]

14. Suppose F and G are two integrable functions on a bounded interval $[a, b]$. Show that G is the weak derivative of F if and only if F can be corrected on a set of measure 0, such that F is absolutely continuous and $F'(x) = G(x)$ for almost every x.

[Hint: If G is the weak derivative of F, use an approximation to show that

$$\int_a^b G(x)\varphi(x)dx = -\int_a^b F(x)\varphi'(x)dx$$

holds for the function φ illustrated in Figure 9.]

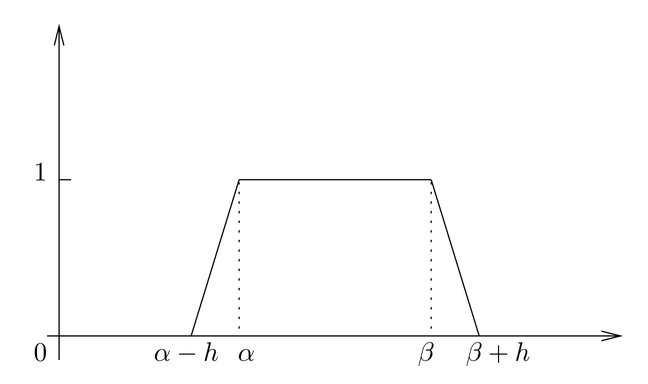


Figure 9. The function φ in Exercise 14

15. Suppose $f \in L^2(\mathbb{R}^d)$. Prove that there exists $g \in L^2(\mathbb{R}^d)$ such that

$$\left(\frac{\partial}{\partial x}\right)^\alpha f(x) = g(x)$$

in the weak sense, if and only if

$$(2\pi i\xi)^\alpha \hat{f}(\xi) = \hat{g}(\xi) \in L^2(\mathbb{R}^d).$$

16. Sobolev embedding theorem. Suppose n is the smallest integer $> d/2$. If

$$f \in L^2(\mathbb{R}^d) \quad \text{and} \quad \left(\frac{\partial}{\partial x}\right)^\alpha f \in L^2(\mathbb{R}^d)$$

in the weak sense, for all $1 \leq |\alpha| \leq n$, then f can be modified on a set of measure zero so that f is continuous and bounded.
[Hint: Express f in terms of $\hat{f}$, and show that $\hat{f} \in L^1(\mathbb{R}^d)$ by the Cauchy-Schwarz inequality.]

17. The conclusion of the Sobolev embedding theorem fails when $n = d/2$. Consider the case $d = 2$, and let $f(x) = (\log 1/|x|)^\alpha \eta(x)$, where η is a smooth cut-off function with $\eta = 1$ for x near the origin, but $\eta(x) = 0$ if $|x| \geq 1/2$. Let $0 < \alpha < 1/2$.

(a) Verify that $\partial f/\partial x_1$ and $\partial f/\partial x_2$ are in L^2 in the weak sense.

(b) Show that f cannot be corrected on a set of measure zero such that the resulting function is continuous at the origin.

18. Consider the linear partial differential operator

$$L = \sum_{|\alpha| \leq n} a_\alpha \left(\frac{\partial}{\partial x}\right)^\alpha.$$

Then

$$P(\xi) = \sum_{|\alpha| \leq n} a_\alpha (2\pi i \xi)^\alpha$$

is called the **characteristic polynomial** of L. The differential operator L is said to be **elliptic** if

$$|P(\xi)| \geq c|\xi|^n \qquad \text{for some } c > 0 \text{ and all } \xi \text{ sufficiently large.}$$

(a) Check that L is elliptic if and only if $\sum_{|\alpha|=n} a_\alpha (2\pi\xi)^\alpha$ vanishes only when $\xi = 0$.

(b) If L is elliptic, prove that for some $c > 0$ the inequality

$$\left\| \left(\frac{\partial}{\partial x} \right)^\alpha \varphi \right\|_{L^2(\mathbb{R}^d)} \leq c \left(\|L\varphi\|_{L^2(\mathbb{R}^d)} + \|\varphi\|_{L^2(\mathbb{R}^d)} \right)$$

holds for all $\varphi \in C_0^\infty(\Omega)$ and $|\alpha| \leq n$.

(c) Conversely, if (b) holds then L is elliptic.

19. Suppose u is harmonic in the punctured unit disc $\mathbb{D}^* = \{z \in \mathbb{C} : 0 < |z| < 1\}$.

(a) Show that if u is also continuous at the origin, then u is harmonic throughout the unit disc.

[Hint: Show that u is weakly harmonic.]

(b) Prove that the Dirichlet problem for the punctured unit disc is in general not solvable.

20. Let F be a continuous function on the closure $\overline{\mathbb{D}}$ of the unit disc. Assume that F is in C^1 on the (open) disc $\mathbb{D}$, and $\int_{\mathbb{D}} |\nabla F|^2 < \infty$.

Let $f(e^{i\theta})$ denote the restriction of F to the unit circle, and write $f(e^{i\theta}) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$. Prove that $\sum_{n=-\infty}^{\infty} |n|\,|a_n|^2 < \infty$.

[Hint: Write $F(re^{i\theta}) \sim \sum_{n=-\infty}^{\infty} F_n(r)e^{in\theta}$, with $F_n(1) = a_n$. Express $\int_{\mathbb{D}} |\nabla F|^2$ in polar coordinates, and use the fact that

$$\frac{1}{2}|F(1)|^2 \leq L^{-1} \int_{1/2}^{1} |F'(r)|^2 \, dr + L \int_{1/2}^{1} |F(r)|^2 \, dr,$$

for $L \geq 2$; apply this to $F = F_n$, $L = |n|$.]

6 Problems

1. Suppose $F_0(x) \in L^2(\mathbb{R})$. Then a necessary and sufficient condition that there exists an entire analytic function F, such that $|F(z)| \leq Ae^{a|z|}$ for all $z \in \mathbb{C}$, and $F_0(x) = F(x)$ a.e. $x \in \mathbb{R}$, is that $\hat{F}_0(\xi) = 0$ whenever $|\xi| > a/2\pi$.

[Hint: Consider the regularization $F^\epsilon(z) = \int_{-\infty}^{\infty} F(z-t)\varphi_\epsilon(t)\,dt$ and apply to it the considerations in Theorem 3.3 of Chapter 4 in Book II.]

2. Suppose Ω is an open bounded subset of $\mathbb{R}^2$. A **boundary Lipschitz arc** γ is a portion of $\partial\Omega$ which after a rotation of the axes is represented as

$$\gamma = \{(x_1, x_2) : x_2 = \eta(x_1),\ a \leq x_1 \leq b\},$$

where $a < b$ and $\gamma \subset \partial\Omega$. It is also supposed that

$$|\eta(x_1) - \eta(x_1')| \leq M|x_1 - x_1'|, \qquad \text{whenever } x_1, x_1' \in [a, b], \tag{35}$$

and moreover if $\gamma_\delta = \{(x_1, x_2) :\ x_2 - \delta \leq \eta(x_1) \leq x_2\}$, then $\gamma_\delta \cap \Omega = \emptyset$ for some $\delta > 0$. (Note that the condition (35) is satisfied if $\eta \in C^1([a, b])$.)

Suppose Ω satisfies the following condition. There are finitely many open discs $D_1, D_2, \ldots, D_N$ with the property that $\bigcup_j D_j$ contains $\partial\Omega$ and for each j, $\partial\Omega \cap D_j$ is a boundary Lipschitz arc (see Figure 10). Then Ω verifies the outside-triangle condition of Theorem 4.12, guaranteeing the solvability of the boundary value problem.

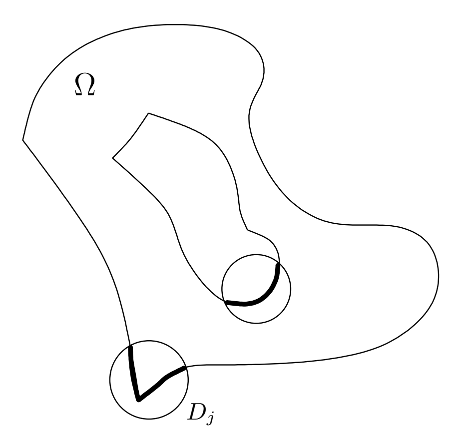


Figure 10. A domain with boundary Lipschitz arcs

3.* Suppose the bounded domain Ω has as its boundary a closed simple continuous curve. Then the boundary value problem is solvable for Ω. This is because there

exists a conformal map Φ of the unit disc $\mathbb{D}$ to Ω that extends to a continuous bijection from $\overline{\mathbb{D}}$ to $\overline{\Omega}$. (See Section 1.3 and Problem 6* in Chapter 8 of Book II.)

4. Consider the two domains Ω in $\mathbb{R}^2$ given by Figure 11.

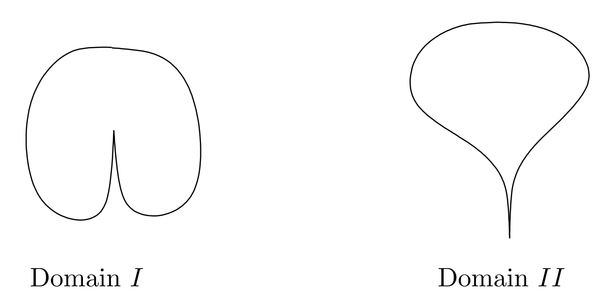


Figure 11. Domains with a cusp

The set I has as its boundary a smooth curve, with the exception of an (inside) cusp. The set II is similar, except it has an outside cusp. Both I and II fall within the scope of the result of Problem 3, and hence the boundary value problem is solvable in each case. However, II satisfies the outside-triangle condition while I does not.

5. Let T be a Fourier multiplier operator on $L^2(\mathbb{R}^d)$. That is, suppose there is a bounded function m such that $\widehat{(Tf)}(\xi) = m(\xi)\hat{f}(\xi)$, all $f \in L^2(\mathbb{R}^d)$. Then T commutes with translations, $\tau_h T = T\tau_h$, where $\tau_h(f)(x) = f(x-h)$, for all $h \in \mathbb{R}^d$.

Conversely any bounded operator on $L^2(\mathbb{R}^d)$ that commutes with translations is a Fourier multiplier operator.

[Hint: It suffices to prove that if a bounded operator $\hat{T}$ commutes with multiplication by exponentials $e^{2\pi i \xi \cdot h}$, $h \in \mathbb{R}^d$, then there is an m so that $\hat{T}g(\xi) = m(\xi)g(\xi)$ for all $g \in L^2(\mathbb{R}^d)$. To do this, show first that

$$\hat{T}(\Phi g) = \Phi\hat{T}(g), \qquad \text{all } g \in L^2(\mathbb{R}^d), \text{ whenever } \Phi \in C_0^\infty(\mathbb{R}^d).$$

Next, for large N, choose Φ so that it equals 1 in the ball $|\xi| \leq N$. Then $m(\xi) = \hat{T}(\Phi)(\xi)$ for $|\xi| \leq N$.]

As a consequence of this theorem show that if T is a bounded operator on $L^2(\mathbb{R})$ that commutes with translations and dilations (as in Exercise 9 above), then

(a) If $(Tf)(-x) = T(f(-x))$ it follows $T = cI$, where c is an appropriate constant and I the identity operator.

(b) If $(Tf)(-x) = -T(f(-x))$, then $T = cH$, where c is an appropriate constant and H the Hilbert transform.

6. This problem provides an example of the contrast between analysis on $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$.

Recall that if f is locally integrable on $\mathbb{R}^d$, the maximal function f^* is defined by

$$f^*(x) = \sup_{x\in B} \frac{1}{m(B)} \int_B |f(y)|\, dy,$$

where the supremum is taken over all balls containing the point x.

Complete the following outline to prove that there exists a constant C so that

$$\|f^*\|_{L^2(\mathbb{R}^d)} \leq C\|f\|_{L^2(\mathbb{R}^d)}.$$

In other words, the map that takes f to f^* (although not linear) is bounded on $L^2(\mathbb{R}^d)$. This differs notably from the situation in $L^1(\mathbb{R}^d)$, as we observed in Chapter 3.

(a) For each $\alpha > 0$, prove that if $f \in L^2(\mathbb{R}^d)$, then

$$m(\{x:\ f^*(x) > \alpha\}) \leq \frac{2A}{\alpha} \int_{|f|>\alpha/2} |f(x)|\, dx.$$

Here, $A = 3^d$ will do.

[Hint: Consider $f_1(x) = f(x)$ if $|f(x)| \geq \alpha/2$ and 0 otherwise. Check that $f_1 \in L^1(\mathbb{R}^d)$, and

$$\{x:\ f^*(x) > \alpha\} \subset \{x:\ f_1^*(x) > \alpha/2\}.]$$

(b) Show that

$$\int_{\mathbb{R}^d} |f^*(x)|^2 dx = 2\int_0^\infty \alpha m(E_\alpha) d\alpha,$$

where $E_\alpha = \{x:\ f^*(x) > \alpha\}$.

(c) Prove that $\|f^*\|_{L^2(\mathbb{R}^d)} \leq C\|f\|_{L^2(\mathbb{R}^d)}$.

6 Abstract Measure and Integration Theory

> What immediately suggest itself, then, is that these characteristic properties themselves be treated as the main object of investigation, by defining and dealing with abstract objects which need satisfy no other conditions than those required by the very theory to be developed.
>
> This procedure has been made use of – more or less consciously – by mathematicians of every era. The geometry of Euclid and the literal algebra of the sixteenth and seventeenth centuries arose in this way. But only in more recent times has this method, called the *axiomatic* method, been consistently developed and carried through to its logical conclusion.
>
> It is our intention to treat the theories of measure and integration by means of the axiomatic method just described.
>
> *C. Carathéodory,* 1918

In much of mathematics integration plays a significant role. It is used, in one form or another, when dealing with questions that arise in analysis on a variety of different spaces. While in some situations it suffices to integrate continuous or other simple functions on these spaces, the deeper study of a number of other problems requires integration based on the more refined ideas of measure theory. The development of these ideas, going beyond the setting of the Euclidean space $\mathbb{R}^d$, is the goal of this chapter.

The starting point is a fruitful insight of Carathéodory and the resulting theorems that lead to construction of measures in very general circumstances. Once this has been achieved, the deduction of the fundamental facts about integration in the general context then follows a familiar path.

We apply the abstract theory to obtain several useful results: the theory of product measures; the polar coordinate integration formula, which is a consequence of this; the construction of the Lebesgue-Stieltjes integral and its corresponding Borel measure on the real line; and the

general notion of absolute continuity. Finally, we treat some of the basic limit theorems of ergodic theory. This not only gives an illustration of the abstract framework we have established, but also provides a link with the differentiation theorems studied in Chapter 3.

1 Abstract measure spaces

A **measure space** consists of a set X equipped with two fundamental objects:

(I) A σ-algebra $\mathcal{M}$ of "measurable" sets, which is a non-empty collection of subsets of X closed under complements and countable unions and intersections.

(II) A **measure** $\mu : \mathcal{M} \to [0, \infty]$ with the following defining property: if $E_1, E_2, \ldots$ is a countable family of disjoint sets in $\mathcal{M}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

A measure space is therefore often denoted by the triple $(X, \mathcal{M}, \mu)$ to emphasize its three main components. Sometimes, however, when there is no ambiguity we will abbreviate this notation by referring to the measure space as (X, μ), or simply X.

A feature that a measure space often enjoys is the property of being **σ-finite**. This means that X can be written as the union of countably many measurable sets of finite measure.

At this early stage we give only two simple examples of measure spaces:

(i) The first is the discrete example with X a countable set, $X = \{x_n\}_{n=1}^{\infty}$, $\mathcal{M}$ the collection of all subsets of X, and the measure μ determined by $\mu(x_n) = \mu_n$, with $\{\mu_n\}_{n=1}^{\infty}$ a given sequence of (extended) non-negative numbers. Note that $\mu(E) = \sum_{x_n \in E} \mu_n$. When $\mu_n = 1$ for all n, we call μ the **counting measure**, and also denote it by $\#$. In this case integration will amount to nothing but the summation of (absolutely) convergent series.

(ii) Here $X = \mathbb{R}^d$, $\mathcal{M}$ is the collection of Lebesgue measurable sets, and $\mu(E) = \int_E f\, dx$, where f is a given non-negative measurable function on $\mathbb{R}^d$. The case $f = 1$ corresponds to the Lebesgue measure. The countable additivity of μ follows from the usual additivity and limiting properties of integrals of non-negative functions proved in Chapter 2.

The construction of measure spaces relevant for most applications require further ideas, and to these we now turn.

1.1 Exterior measures and Carathéodory's theorem

To begin the construction of a measure and its corresponding measurable sets in the general setting requires, as in the special case of Lebesgue measure considered in Chapter 1, a prerequisite notion of "exterior" measure. This is defined as follows.

Let X be a set. An **exterior measure** (or **outer measure**) μ_* on X is a function μ_* from the collection of *all* subsets of X to $[0, \infty]$ that satisfies the following properties:

(i) $\mu_*(\emptyset) = 0$.

(ii) If $E_1 \subset E_2$, then $\mu_*(E_1) \leq \mu_*(E_2)$.

(iii) If $E_1, E_2, \ldots$ is a countable family of sets, then

$$\mu_*\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu_*(E_j).$$

For instance, the exterior Lebesgue measure m_* in $\mathbb{R}^d$ defined in Chapter 1 enjoys all these properties. In fact, this example belongs to a large class of exterior measures that can be obtained using "coverings" by a family of special sets whose measures are taken as known. This idea is systematized by the notion of a "premeasure" taken up below in Section 1.3. A different type of example is the exterior α-dimensional Hausdorff measure m_α^* defined in Chapter 7.

Given an exterior measure μ_*, the problem that one faces is how to define the corresponding notion of measurable sets. In the case of Lebesgue measure in $\mathbb{R}^d$ such sets were characterized by their difference from open (or closed) sets, when considered in terms of μ_*. For the general case, Carathéodory found an ingenious substitute condition. It is as follows.

A set E in X is **Carathéodory measurable** or simply **measurable** if one has

$$\mu_*(A) = \mu_*(E \cap A) + \mu_*(E^c \cap A) \quad \text{for every } A \subset X. \tag{1}$$

In other words, E separates any set A in two parts that behave well in regard to the exterior measure μ_*. For this reason, (1) is sometimes referred to as the separation condition. One can show that in $\mathbb{R}^d$ with the Lebesgue exterior measure the notion of measurability (1) is equivalent

to the definition of Lebesgue measurability given in Chapter 1. (See Exercise 3.)

A first observation we make is that to prove a set E is measurable, it suffices to verify

$$\mu_*(A) \geq \mu_*(E \cap A) + \mu_*(E^c \cap A) \quad \text{for all } A \subset X,$$

since the reverse inequality is automatically verified by the sub-additivity property (iii) of the exterior measure. We see immediately from the definition that sets of exterior measure zero are necessarily measurable.

The remarkable fact about the definition (1) is summarized in the next theorem.

Theorem 1.1 *Given an exterior measure μ_* on a set X, the collection $\mathcal{M}$ of Carathéodory measurable sets forms a σ-algebra. Moreover, μ_* restricted to $\mathcal{M}$ is a measure.*

Proof. Clearly, $\emptyset$ and X belong to $\mathcal{M}$ and the symmetry inherent in condition (1) shows that $E^c \in \mathcal{M}$ whenever $E \in \mathcal{M}$. Thus $\mathcal{M}$ is non-empty and closed under complements.

Next, we prove that $\mathcal{M}$ is closed under finite unions of disjoint sets, and μ_* is finitely additive on $\mathcal{M}$. Indeed, if $E_1, E_2 \in \mathcal{M}$, and A is any subset of X, then

$$\begin{aligned}
\mu_*(A) &= \mu_*(E_2 \cap A) + \mu_*(E_2^c \cap A) \\
&= \mu_*(E_1 \cap E_2 \cap A) + \mu_*(E_1^c \cap E_2 \cap A) + \\
&\qquad\qquad + \mu_*(E_1 \cap E_2^c \cap A) + \mu_*(E_1^c \cap E_2^c \cap A) \\
&\geq \mu_*((E_1 \cup E_2) \cap A) + \mu_*((E_1 \cup E_2)^c \cap A),
\end{aligned}$$

where in the first two lines we have used the measurability condition on E_2 and then E_1, and where the last inequality was obtained using the sub-additivity of μ_* and the fact that $E_1 \cup E_2 = (E_1 \cap E_2) \cup (E_1^c \cap E_2) \cup (E_1 \cap E_2^c)$. Therefore, we have $E_1 \cup E_2 \in \mathcal{M}$, and if E_1 and E_2 are disjoint, we find

$$\begin{aligned}
\mu_*(E_1 \cup E_2) &= \mu_*\left(E_1 \cap (E_1 \cup E_2)\right) + \mu_*\left(E_1^c \cap (E_1 \cup E_2)\right) \\
&= \mu_*(E_1) + \mu_*(E_2).
\end{aligned}$$

Finally, it suffices to show that $\mathcal{M}$ is closed under countable unions of disjoint sets, and that μ_* is countably additive on $\mathcal{M}$. Let $E_1, E_2, \ldots$ denote a countable collection of disjoint sets in $\mathcal{M}$, and define

$$G_n = \bigcup_{j=1}^{n} E_j \quad \text{and} \quad G = \bigcup_{j=1}^{\infty} E_j.$$

For each n, the set G_n is a finite union of sets in $\mathcal{M}$, hence $G_n \in \mathcal{M}$. Moreover, for any $A \subset X$ we have

$$\begin{aligned}
\mu_*(G_n \cap A) &= \mu_*(E_n \cap (G_n \cap A)) + \mu_*(E_n^c \cap (G_n \cap A)) \\
&= \mu_*(E_n \cap A) + \mu_*(G_{n-1} \cap A) \\
&= \sum_{j=1}^{n} \mu_*(E_j \cap A),
\end{aligned}$$

where the last equality is obtained by induction. Since we know that $G_n \in \mathcal{M}$, and $G^c \subset G_n^c$, we find that

$$\mu_*(A) = \mu_*(G_n \cap A) + \mu_*(G_n^c \cap A) \geq \sum_{j=1}^{n} \mu_*(E_j \cap A) + \mu_*(G^c \cap A).$$

Letting n tend to infinity, we obtain

$$\begin{aligned}
\mu_*(A) \geq \sum_{j=1}^{\infty} \mu_*(E_j \cap A) + \mu_*(G^c \cap A) &\geq \mu_*(G \cap A) + \mu_*(G^c \cap A) \\
&\geq \mu_*(A).
\end{aligned}$$

Therefore all the inequalities above are equalities, and we conclude that $G \in \mathcal{M}$, as desired. Moreover, by taking $A = G$ in the above, we find that μ_* is countably additive on $\mathcal{M}$, and the proof of the theorem is complete.

Our previous observation that sets of exterior measure 0 are Carathéodory measurable shows that the measure space $(X, \mathcal{M}, \mu)$ in the theorem is **complete**: whenever $F \in \mathcal{M}$ satisfies $\mu(F) = 0$ and $E \subset F$, then $E \in \mathcal{M}$.

1.2 Metric exterior measures

If the underlying set X is endowed with a "distance function" or "metric," there is a particular class of exterior measures that is of interest in practice. The importance of these exterior measures is that they induce measures on the natural σ-algebra generated by the open sets in X.

A **metric space** is a set X equipped with a function $d : X \times X \to [0, \infty)$ that satisfies:

(i) $d(x, y) = 0$ if and only if $x = y$.

(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$.

(iii) $d(x, z) \leq d(x, y) + d(y, z)$, for all $x, y, z \in X$.

The last property is of course called the triangle inequality, and a function d that satisfies all these conditions is called a **metric** on X. For example, the set $\mathbb{R}^d$ with $d(x, y) = |x - y|$ is a metric space. Another example is provided by the space of continuous functions on a compact set K with $d(f, g) = \sup_{x \in K} |f(x) - g(x)|$.

A metric space (X, d) is naturally equipped with a family of open balls. Here

$$B_r(x) = \{y \in X : d(x, y) < r\}$$

defines the open ball of radius r centered at x. Together with this, we say that a set $\mathcal{O} \subset X$ is **open** if for any $x \in \mathcal{O}$ there exists $r > 0$ so that the open ball $B_r(x)$ is contained in $\mathcal{O}$. A set is **closed** if its complement is open. With these definitions, one checks easily that an (arbitrary) union of open sets is open, and a similar intersection of closed sets is closed.

Finally, on a metric space X we can define, as in Section 3 of Chapter 1, the **Borel σ-algebra**, $\mathcal{B}_X$, that is the smallest σ-algebra of sets in X that contains the open sets of X. In other words $\mathcal{B}_X$ is the intersection of all σ-algebras that contain the open sets. Elements in $\mathcal{B}_X$ are called **Borel sets**.

We now turn our attention to those exterior measures on X with the special property of being additive on sets that are "well separated." We show that this property guarantees that this exterior measure defines a measure on the Borel σ-algebra. This is achieved by proving that all Borel sets are Carathéodory measurable.

Given two sets A and B in a metric space (X, d), the **distance** between A and B is defined by

$$d(A, B) = \inf\{d(x, y) : \ x \in A \text{ and } y \in B\}.$$

Then an exterior measure μ_* on X is a **metric exterior measure** if it satisfies

$$\mu_*(A \cup B) = \mu_*(A) + \mu_*(B) \qquad \text{whenever } d(A, B) > 0.$$

This property played a key role in the case of exterior Lebesgue measure.

Theorem 1.2 *If μ_* is a metric exterior measure on a metric space X, then the Borel sets in X are measurable. Hence μ_* restricted to $\mathcal{B}_X$ is a measure.*

Proof. By the definition of $\mathcal{B}_X$ it suffices to prove that closed sets in X are Carathéodory measurable. Therefore, let F denote a closed set and A a subset of X with $\mu_*(A) < \infty$. For each $n > 0$, let

$$A_n = \{x \in F^c \cap A : \quad d(x, F) \geq 1/n\}.$$

Then $A_n \subset A_{n+1}$, and since F is closed we have $F^c \cap A = \bigcup_{n=1}^{\infty} A_n$. Also, the distance between $F \cap A$ and A_n is $\geq 1/n$, and since μ_* is a metric exterior measure, we have

$$\mu_*(A) \geq \mu_*((F \cap A) \cup A_n) = \mu_*(F \cap A) + \mu_*(A_n). \tag{2}$$

Next, we claim that

$$\lim_{n \to \infty} \mu_*(A_n) = \mu_*(F^c \cap A). \tag{3}$$

To see this, let $B_n = A_{n+1} \cap A_n^c$ and note that

$$d(B_{n+1}, A_n) \geq \frac{1}{n(n+1)}.$$

Indeed, if $x \in B_{n+1}$ and $d(x, y) < 1/n(n+1)$ the triangle inequality shows that $d(y, F) < 1/n$, hence $y \notin A_n$. Therefore

$$\mu_*(A_{2k+1}) \geq \mu_*(B_{2k} \cup A_{2k-1}) = \mu_*(B_{2k}) + \mu_*(A_{2k-1}),$$

and this implies that

$$\mu_*(A_{2k+1}) \geq \sum_{j=1}^{k} \mu_*(B_{2j}).$$

A similar argument also gives

$$\mu_*(A_{2k}) \geq \sum_{j=1}^{k} \mu_*(B_{2j-1}).$$

Since $\mu_*(A)$ is finite, we find that both series $\sum \mu_*(B_{2j})$ and $\sum \mu_*(B_{2j-1})$ are convergent. Finally, we note that

$$\mu_*(A_n) \leq \mu_*(F^c \cap A) \leq \mu_*(A_n) + \sum_{j=n+1}^{\infty} \mu_*(B_j),$$

and this proves the limit (3). Letting n tend to infinity in the inequality (2) we find that $\mu_*(A) \geq \mu_*(F \cap A) + \mu_*(F^c \cap A)$, and hence F is measurable, as was to be shown.

Given a metric space X, a measure μ defined on the Borel sets of X will be referred to as a **Borel measure**. Borel measures that assign a finite measure to all balls (of finite radius) also satisfy a useful regularity property. The requirement that $\mu(B) < \infty$ for all balls B is satisfied in many (but not in all) circumstances that arise in practice.[1] When it does hold, we get the following proposition.

Proposition 1.3 *Suppose the Borel measure μ is finite on all balls in X of finite radius. Then for any Borel set E and any $\epsilon > 0$, there are an open set $\mathcal{O}$ and a closed set F such that $E \subset \mathcal{O}$ and $\mu(\mathcal{O} - E) < \epsilon$, while $F \subset E$ and $\mu(E - F) < \epsilon$.*

Proof. We need the following preliminary observation. Suppose $F^* = \bigcup_{k=1}^{\infty} F_k$, where the F_k are closed sets. Then for any $\epsilon > 0$, we can find a closed set $F \subset F^*$ such that $\mu(F^* - F) < \epsilon$. To prove this we can assume that the sets $\{F_k\}$ are increasing. Fix a point $x_0 \in X$, and let B_n denote the ball $\{x : d(x, x_0) < n\}$, with $B_0 = \{\emptyset\}$. Since $\bigcup_{n=1}^{\infty} B_n = X$, we have that

$$F^* = \bigcup F^* \cap (\overline{B}_n - B_{n-1}).$$

Now for each n, $F^* \cap (\overline{B}_n - B_{n-1})$ is the limit as $k \to \infty$ of the increasing sequence of closed sets $F_k \cap (\overline{B}_n - B_{n-1})$, so (recalling that $\overline{B_n}$ has finite measure) we can find an $N = N(n)$ so that $(F^* - F_{N(n)}) \cap (\overline{B}_n - B_{n-1})$ has measure less than $\epsilon/2^n$. If we now let

$$F = \bigcup_{n=1}^{\infty} \left(F_{N(n)} \cap (\overline{B}_n - B_{n-1}) \right),$$

it follows that the measure of $F^* - F$ is less that $\sum_{n=1}^{\infty} \epsilon/2^n = \epsilon$. We also see that $F \cap \overline{B}_k$ is closed since it is the finite union of closed sets. Thus F itself is closed because, as is easily seen, any set F is closed whenever the sets $F \cap \overline{B}_k$ are closed for all k.

Having established the observation, we call $\mathcal{C}$ the collection of all sets that satisfy the conclusions of the proposition. Notice first that if E belongs to $\mathcal{C}$ then automatically so does its complement.

[1] This restriction is not always valid for the Hausdorff measures that are considered in the next chapter.

Suppose now that $E = \bigcup_{k=1}^\infty E_k$, with each $E_k \in \mathcal{C}$. Then there are open sets $\mathcal{O}_k$, $\mathcal{O}_k \supset E_k$, with $\mu(\mathcal{O}_k - E_k) < \epsilon/2^k$. However, if $\mathcal{O} = \bigcup_{k=1}^\infty \mathcal{O}_k$, then $\mathcal{O} - E \subset \bigcup_{k=1}^\infty (\mathcal{O}_k - E_k)$, and so $\mu(\mathcal{O} - E) \leq \sum_{k=1}^\infty \epsilon/2^k = \epsilon$.

Next, there are closed sets $F_k \subset E_k$ with $\mu(E_k - F_k) < \epsilon/2^k$. Thus if $F^* = \bigcup_{k=1}^\infty F_k$, we see as before that $\mu(E - F^*) < \epsilon$. However, F^* is not necessarily closed, so we can use our preliminary observation to find a closed set $F \subset F^*$ with $\mu(F^* - F) < \epsilon$. Thus $\mu(E - F) < 2\epsilon$. Since ϵ is arbitrary, this proves that $\bigcup_{k=1}^\infty E_k$ belongs to $\mathcal{C}$.

Let us finally note that any open set $\mathcal{O}$ is in $\mathcal{C}$. The property regarding containment by open sets is immediate. To find a closed $F \subset \mathcal{O}$, so that $\mu(\mathcal{O} - F) < \epsilon$, let $F_k = \{x \in \overline{B}_k : \ d(x, \mathcal{O}^c) \geq 1/k\}$. Then it is clear that each F_k is closed and $\mathcal{O} = \bigcup_{k=1}^\infty F_k$. We then need only apply the observation again to find the required set F. Thus we have shown that $\mathcal{C}$ is a σ-algebra that contains the open sets, and hence all Borel sets. The proposition is therefore proved.

1.3 The extension theorem

As we have seen, a class of measurable sets on X can be constructed once we start with a given exterior measure. However, the definition of an exterior measure usually depends on a more primitive idea of measure defined on a simpler class of sets. This is the role of a premeasure defined below. As we will show, any premeasure can be extended to a measure on X. We begin with several definitions.

Let X be a set. An **algebra** in X is a non-empty collection of subsets of X that is closed under complements, *finite* unions, and *finite* intersections. Let $\mathcal{A}$ be an algebra in X. A **premeasure** on an algebra $\mathcal{A}$ is a function $\mu_0 : \mathcal{A} \to [0, \infty]$ that satisfies:

(i) $\mu_0(\emptyset) = 0$.

(ii) If $E_1, E_2, \ldots$ is a countable collection of disjoint sets in $\mathcal{A}$ with $\bigcup_{k=1}^\infty E_k \in \mathcal{A}$, then

$$\mu_0\left(\bigcup_{k=1}^\infty E_k\right) = \sum_{k=1}^\infty \mu_0(E_k).$$

In particular, μ_0 is finitely additive on $\mathcal{A}$.

Premeasures give rise to exterior measures in a natural way.

Lemma 1.4 *If μ_0 is a premeasure on an algebra $\mathcal{A}$, define μ_* on any subset E of X by*

$$\mu_*(E) = \inf\left\{\sum_{j=1}^{\infty}\mu_0(E_j) : \; E \subset \bigcup_{j=1}^{\infty} E_j, \; \textit{where } E_j \in \mathcal{A} \textit{ for all } j\right\}.$$

Then, μ_* is an exterior measure on X that satisfies:

(i) *$\mu_*(E) = \mu_0(E)$ for all $E \in \mathcal{A}$.*

(ii) *All sets in $\mathcal{A}$ are measurable in the sense of* (1).

Proof. Proving that μ_* is an exterior measure presents no difficulty. To see why the restriction of μ_* to $\mathcal{A}$ coincides with μ_0, suppose that $E \in \mathcal{A}$. Clearly, one always has $\mu_*(E) \leq \mu_0(E)$ since E covers itself. To prove the reverse inequality let $E \subset \bigcup_{j=1}^{\infty} E_j$, where $E_j \in \mathcal{A}$ for all j. Then, if we set

$$E_k' = E \cap \left(E_k - \bigcup_{j=1}^{k-1} E_j\right),$$

the sets E_k' are disjoint elements of $\mathcal{A}$, $E_k' \subset E_k$ and $E = \bigcup_{k=1}^{\infty} E_k'$. By (ii) in the definition of a premeasure, we have

$$\mu_0(E) = \sum_{k=1}^{\infty}\mu_0(E_k') \leq \sum_{k=1}^{\infty}\mu_0(E_k).$$

Therefore, we find that $\mu_0(E) \leq \mu_*(E)$, as desired.

Finally, we must prove that sets in $\mathcal{A}$ are measurable for μ_*. Let A be any subset of X, $E \in \mathcal{A}$, and $\epsilon > 0$. By definition, there exists a countable collection $E_1, E_2, \ldots$ of sets in $\mathcal{A}$ such that $A \subset \bigcup_{j=1}^{\infty} E_j$ and

$$\sum_{j=1}^{\infty}\mu_0(E_j) \leq \mu_*(A) + \epsilon.$$

Since μ_0 is a premeasure, it is finitely additive on $\mathcal{A}$ and therefore

$$\begin{aligned}\sum_{j=1}^{\infty}\mu_0(E_j) &= \sum_{j=1}^{\infty}\mu_0(E \cap E_j) + \sum_{j=1}^{\infty}\mu_0(E^c \cap E_j)\\ &\geq \mu_*(E \cap A) + \mu_*(E^c \cap A).\end{aligned}$$

Since ϵ is arbitrary, we conclude that $\mu_*(A) \geq \mu_*(E \cap A) + \mu_*(E^c \cap A)$, as desired.

The σ-algebra generated by an algebra $\mathcal{A}$ is by definition the smallest σ-algebra that contains $\mathcal{A}$. The above lemma then provides the necessary step for extending μ_0 on $\mathcal{A}$ to a measure on the σ-algebra generated by $\mathcal{A}$.

Theorem 1.5 *Suppose that $\mathcal{A}$ is an algebra of sets in X, μ_0 a premeasure on $\mathcal{A}$, and $\mathcal{M}$ the σ-algebra generated by $\mathcal{A}$. Then there exists a measure μ on $\mathcal{M}$ that extends μ_0.*

One notes below that μ is the only such extension of μ_0 under the assumption that μ is σ-finite.

Proof. The exterior measure μ_* induced by μ_0 defines a measure μ on the σ-algebra of Carathéodory measurable sets. Therefore, by the result in the previous lemma, μ is also a measure on $\mathcal{M}$ that extends μ_0. (We should observe that in general the class $\mathcal{M}$ is not as large as the class of all sets that are measurable in the sense of (1).)

To prove that this extension is unique whenever μ is σ-finite, we argue as follows. Suppose that ν is another measure on $\mathcal{M}$ that coincides with μ_0 on $\mathcal{A}$, and suppose that $F \in \mathcal{M}$ has finite measure. We claim that $\mu(F) = \nu(F)$. If $F \subset \bigcup E_j$, where $E_j \in \mathcal{A}$, then

$$\nu(F) \leq \sum_{j=1}^{\infty} \nu(E_j) = \sum_{j=1}^{\infty} \mu_0(E_j),$$

so that $\nu(F) \leq \mu(F)$. To prove the reverse inequality, note that if $E = \bigcup E_j$, then the fact that ν and μ are two measures that agree on $\mathcal{A}$ gives

$$\nu(E) = \lim_{n\to\infty} \nu(\bigcup_{j=1}^{n} E_j) = \lim_{n\to\infty} \mu(\bigcup_{j=1}^{n} E_j) = \mu(E).$$

If the sets E_j are chosen so that $\mu(E) \leq \mu(F) + \epsilon$, then the fact that $\mu(F) < \infty$ implies $\mu(E - F) \leq \epsilon$, and therefore

$$\begin{aligned}\mu(F) \leq \mu(E) = \nu(E) = \nu(F) + \nu(E - F) &\leq \nu(F) + \mu(E - F) \\ &\leq \mu(F) + \epsilon.\end{aligned}$$

Since ϵ is arbitrary, we find that $\mu(F) \leq \nu(F)$, as desired.

Finally, we use this last result to prove that if μ is σ-finite, then $\mu = \nu$. Indeed, we may write $X = \bigcup E_j$, where $E_1, E_2, \ldots$ is a countable

collection of disjoint sets in $\mathcal{A}$ with $\mu(E_j) < \infty$. Then for any $F \in \mathcal{M}$ we have

$$\mu(F) = \sum \mu(F \cap E_j) = \sum \nu(F \cap E_j) = \nu(F),$$

and the uniqueness is proved.

For later use we record the following observation about the premeasure μ_0 on the algebra $\mathcal{A}$ and the resulting measure μ_* that is implicit in the argument given above. The details of the proof may be left to the reader.

We define $\mathcal{A}_\sigma$ as the collection of sets that are countable unions of sets in $\mathcal{A}$, and $\mathcal{A}_{\sigma\delta}$ as the sets that arise as countable intersections of sets in $\mathcal{A}_\sigma$.

Proposition 1.6 *For any set E and any $\epsilon > 0$, there are sets $E_1 \in \mathcal{A}_\sigma$ and $E_2 \in \mathcal{A}_{\sigma\delta}$, such that $E \subset E_1$, $E \subset E_2$, and $\mu_*(E_1) \leq \mu_*(E) + \epsilon$, while $\mu_*(E_2) = \mu_*(E)$.*

2 Integration on a measure space

Once we have established the basic properties of a measure space X, the fundamental facts about measurable functions and integration of such functions on X can be deduced as in the case of the Lebesgue measure on $\mathbb{R}^d$. Indeed, the results in Section 4 of Chapter 1 and all of Chapter 2 go over to the general case, with proofs remaining almost word-for-word the same. For this reason we shall not repeat these arguments but limit ourselves to the bare statement of the main points. The reader should have no difficulty in filling in the missing details.

To avoid unnecessary complications we will assume throughout that the measure space $(X, \mathcal{M}, \mu)$ under consideration is σ-finite.

Measurable functions

A function f on X with values in the extended real numbers is **measurable** if

$$f^{-1}([-\infty, a)) = \{x \in X : f(x) < a\} \in \mathcal{M} \qquad \text{for all } a \in \mathbb{R}.$$

With this definition, the basic properties of measurable functions obtained in the case of $\mathbb{R}^d$ with the Lebesgue measure continue to hold. (See Properties 3 through 6 for measurable functions in Chapter 1.) For instance, the collection of measurable functions is closed under the basic algebraic manipulations. Also, the pointwise limits of measurable functions are measurable.

The notion of "almost everywhere" that we use now is with respect to the measure μ. For instance, if f and g are measurable functions on X, we write $f = g$ a.e. to say that

$$\mu(\{x \in X : f(x) \neq g(x)\}) = 0.$$

A **simple function** on X takes the form

$$\sum_{k=1}^{N} a_k \chi_{E_k},$$

where E_k are measurable sets of finite measure and a_k are real numbers. Approximations by simple functions played an important role in the definition of the Lebesgue integral. Fortunately, this result continues to hold in our abstract setting.

- *Suppose f is a non-negative measurable function on a measure space $(X, \mathcal{M}, \mu)$. Then there exists a sequence of simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that satisfies*

  $$\varphi_k(x) \leq \varphi_{k+1}(x) \quad \textit{and} \quad \lim_{k\to\infty} \varphi_k(x) = f(x) \textit{ for all } x.$$

  In general, if f is only measurable, there exists a sequence of simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that satisfies

  $$|\varphi_k(x)| \leq |\varphi_{k+1}(x)| \quad \textit{and} \quad \lim_{k\to\infty} \varphi_k(x) = f(x) \textit{ for all } x.$$

The proof of this result can be obtained with some obvious minor modifications of the proofs of Theorems 4.1 and 4.2 in Chapter 1. Here, one makes use of the technical condition imposed on X, that of being σ-finite. Indeed, if we write $X = \bigcup F_k$, where $F_k \in \mathcal{M}$ are of finite measure, then the sets F_k play the role of the cubes Q_k in the proof of Theorem 4.1, Chapter 1.

Another important result that generalizes immediately is Egorov's theorem.

- *Suppose $\{f_k\}_{k=1}^{\infty}$ is a sequence of measurable functions defined on a measurable set $E \subset X$ with $\mu(E) < \infty$, and $f_k \to f$ a.e. Then for each $\epsilon > 0$ there is a set A_ϵ with $A_\epsilon \subset E$, $\mu(E - A_\epsilon) \leq \epsilon$, and such that $f_k \to f$ uniformly on A_ϵ.*

Definition and main properties of the integral

The four-step approach to the construction of the Lebesgue integral that begins with its definition on simple functions given in Chapter 2 carries over to the situation of a σ-finite measure space $(X, \mathcal{M}, \mu)$. This leads to the notion of the integral, with respect to the measure μ, of a non-negative measurable function f on X. This integral is denoted by

$$\int_X f(x)\, d\mu(x),$$

which we sometimes simplify as $\int_X f\, d\mu$, $\int f\, d\mu$ or $\int f$, when no confusion is possible. Finally, we say that a measurable function f is **integrable** if

$$\int_X |f(x)|\, d\mu(x) < \infty.$$

The elementary properties of the integral, such as linearity and monotonicity, continue to hold in this general setting, as well as the following basic limit theorems.

(i) Fatou's lemma. *If $\{f_n\}$ is a sequence of non-negative measurable functions on X, then*

$$\int \liminf_{n\to\infty} f_n\, d\mu \leq \liminf_{n\to\infty} \int f_n\, d\mu.$$

(ii) Monotone convergence. *If $\{f_n\}$ is a sequence of non-negative measurable functions with $f_n \nearrow f$, then*

$$\lim_{n\to\infty} \int f_n = \int f.$$

(iii) Dominated convergence. *If $\{f_n\}$ is a sequence of measurable functions with $f_n \to f$ a.e., and such that $|f_n| \leq g$ for some integrable g, then*

$$\int |f_n - f|\, d\mu \to 0 \qquad \textit{as } n \to \infty,$$

and consequently

$$\int f_n\, d\mu \to \int f\, d\mu \qquad \textit{as } n \to \infty.$$

The spaces $L^1(X,\mu)$ and $L^2(X,\mu)$

The equivalence classes (modulo functions that vanish almost everywhere) of integrable functions on $(X, \mathcal{M}, \mu)$ form a vector space equipped with a norm. This space is denoted by $L^1(X,\mu)$ and its norm is

$$\|f\|_{L^1(X,\mu)} = \int_X |f(x)|\, d\mu(x). \tag{4}$$

Similarly we can define $L^2(X,\mu)$ to be the equivalence class of measurable functions for which $\int_X |f(x)|^2\, d\mu(x) < \infty$. The norm is then

$$\|f\|_{L^2(X,\mu)} = \left(\int_X |f(x)|^2\, d\mu(x)\right)^{1/2}. \tag{5}$$

There is also an inner product on this space given by

$$(f, g) = \int_X f(x)\overline{g(x)}\, d\mu(x).$$

The proofs of Proposition 2.1 and Theorem 2.2 in Chapter 2, as well as the results in Section 1 of Chapter 4, extend to this general case and give:

- *The space $L^1(X,\mu)$ is a complete normed vector space.*
- *The space $L^2(X,\mu)$ is a (possibly non-separable) Hilbert space.*

3 Examples

We now discuss some useful examples of the general theory.

3.1 Product measures and a general Fubini theorem

Our first example concerns the construction of product measures, and leads to a general form of the theorem that expresses a multiple integral as a repeated integral, extending the case of Euclidean space considered in Section 3 of Chapter 2.

Suppose $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ are a pair of measure spaces. We want to describe the **product measure** $\mu_1 \times \mu_2$ on the space $X = X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1,\ x_2 \in X_2\}$.

We will assume here that the two measure spaces are each complete and σ-finite.

We begin by considering **measurable rectangles**: these are subsets of X of the form $A \times B$, with A and B measurable sets, that is, $A \in \mathcal{M}_1$

and $B \in \mathcal{M}_2$. We then let $\mathcal{A}$ denote the collection of all sets in X that are finite unions of disjoint measurable rectangles. It is easy to check that $\mathcal{A}$ is an algebra of subsets of X. (Indeed, the complement of a measurable rectangle is the union of three disjoint such rectangles, while the union of two measurable rectangles is the disjoint union of at most six such rectangles.) From now on we abbreviate our terminology by referring to measurable rectangles simply as "rectangles."

On the rectangles we define the function μ_0 by $\mu_0(A \times B) = \mu_1(A)\mu_2(B)$. Now the fact that μ_0 has a unique extension to the algebra $\mathcal{A}$ for which μ_0 becomes a premeasure is a consequence of the following fact: whenever a rectangle $A \times B$ is the disjoint union of a countable collection of rectangles $\{A_j \times B_j\}$, $A \times B = \bigcup_{j=1}^\infty A_j \times B_j$, then

$$\mu_0(A \times B) = \sum_{j=1}^{\infty} \mu_0(A_j \times B_j). \tag{6}$$

To prove this, observe that if $x_1 \in A$, then for each $x_2 \in B$ the point (x_1, x_2) belongs to exactly one $A_j \times B_j$. Therefore we see that B is the disjoint union of the B_j for which $x_1 \in A_j$. By the countable additivity property of the measure μ_2 this has as an immediate consequence the fact that

$$\chi_A(x_1)\mu_2(B) = \sum_{j=1}^{\infty} \chi_{A_j}(x_1)\mu_2(B_j).$$

Hence integrating in x_1 and using the monotone convergence theorem we get $\mu_1(A)\mu_2(B) = \sum_{j=1}^\infty \mu_1(A_j)\mu_2(B_j)$, which is (6).

Now that we know that μ_0 is a premeasure on $\mathcal{A}$, we obtain from Theorem 1.5 a measure (which we denote by $\mu = \mu_1 \times \mu_2$) on the σ-algebra $\mathcal{M}$ of sets generated by the algebra $\mathcal{A}$ of measurable rectangles. In this way, we have defined the product measure space $(X_1 \times X_2, \mathcal{M}, \mu_1 \times \mu_2)$.

Given a set E in $\mathcal{M}$ we shall now consider slices

$$E_{x_1} = \{x_2 \in X_2 : (x_1, x_2) \in E\} \quad \text{and} \quad E^{x_2} = \{x_1 \in X_1 : (x_1, x_2) \in E\}.$$

We recall the definitions according to which $\mathcal{A}_\sigma$ denotes the collection of sets that are countable unions of elements of $\mathcal{A}$, and $\mathcal{A}_{\sigma\delta}$ the sets that arise as countable intersections of sets from $\mathcal{A}_\sigma$. We then have the following key fact.

Proposition 3.1 *If E belongs to $\mathcal{A}_{\sigma\delta}$, then E^{x_2} is μ_1-measurable for every x_2; moreover, $\mu_1(E^{x_2})$ is a μ_2-measurable function. In addition*

$$\int_{X_2} \mu_1(E^{x_2})\, d\mu_2 = (\mu_1 \times \mu_2)(E). \tag{7}$$

Proof. One notes first that all the assertions hold immediately when E is a (measurable) rectangle. Next suppose E is a set in $\mathcal{A}_\sigma$. Then we can decompose it as a countable union of *disjoint* rectangles E_j. (If the E_j are not already disjoint we only need to replace the E_j by $\bigcup_{k\le j} E_k - \bigcup_{k\le j-1} E_k$.) Then for each x_2 we have $E^{x_2} = \bigcup_{j=1}^\infty E_j^{x_2}$, and we observe that $\{E_j^{x_2}\}$ are disjoint sets. Thus by (7) applied to each rectangle E_j and the monotone convergence theorem we get our conclusion for each set $E \in \mathcal{A}_\sigma$.

Next assume $E \in \mathcal{A}_{\sigma\delta}$ and that $(\mu_1 \times \mu_2)(E) < \infty$. Then there is a sequence $\{E_j\}$ of sets with $E_j \in \mathcal{A}_\sigma$, $E_{j+1} \subset E_j$, and $E = \bigcap_{j=1}^\infty E_j$. We let $f_j(x_2) = \mu_1(E_j^{x_2})$ and $f(x_2) = \mu_1(E^{x_2})$. To see that E^{x_2} is μ_1-measurable and $f(x_2)$ is well-defined, note that E^{x_2} is the decreasing limit of the sets $E_j^{x_2}$, which we have seen by the above are measurable. Moreover, since $E_1 \in \mathcal{A}_\sigma$ and $(\mu_1 \times \mu_2)(E_1) < \infty$, we see that $f_j(x_2) \to f(x_2)$, as $j \to \infty$ for each x_2. Thus $f(x_2)$ is measurable. However, $\{f_j(x_2)\}$ is a decreasing sequence of non-negative functions, hence

$$\int_{X_2} f(x_2)\, d\mu_2(x) = \lim_{j\to\infty} \int_{X_2} f_j(x_2)\, d\mu_2(x),$$

and therefore (7) is proved in the case when $(\mu_1 \times \mu_2)(E) < \infty$. Now since we assumed both μ_1 and μ_2 are σ-finite, we can find sequences $F_1 \subset F_2 \subset \cdots \subset F_j \subset \cdots \subset X_1$ and $G_1 \subset G_2 \subset \cdots \subset G_j \subset \cdots \subset X_2$, with $\bigcup_{j=1}^\infty F_j = X_1$, $\bigcup_{j=1}^\infty G_j = X_2$, $\mu_1(F_j) < \infty$, and $\mu_2(G_j) < \infty$ for all j. Then we merely need to replace E by $E_j = E \cap (F_j \times G_j)$, and let $j \to \infty$ to obtain the general result.

We now extend the result in the above proposition to an arbitrary measurable set E in $X_1 \times X_2$, that is, $E \in \mathcal{M}$, the σ-algebra generated by the measurable rectangles.

Proposition 3.2 *If E is an arbitrary measurable set in X, then the conclusion of Proposition 3.1 are still valid except that we only assert that E^{x_2} is μ_1-measurable and $\mu_1(E^{x_2})$ is defined for almost every $x_2 \in X_2$.*

Proof. Consider first the case when E is a set of measure zero. Then we know by Proposition 1.6 that there is a set $F \in \mathcal{A}_{\sigma\delta}$ such that

$E \subset F$ and $(\mu_1 \times \mu_2)(F) = 0$. Since $E^{x_2} \subset F^{x_2}$ for every x_2 and F^{x_2} has μ_1-measure zero for almost every x_2 by (7) applied to F, the assumed completeness of the measure μ_2 shows that E^{x_2} is measurable and has measure zero for those x_2. Thus the desired conclusion holds when E has measure zero.

If we drop this assumption on E, we can invoke Proposition 1.6 again to find an $F \in \mathcal{A}_{\sigma\delta}$, $F \supset E$, such that $F - E = Z$ has measure zero. Since $F^{x_2} - E^{x_2} = Z^{x_2}$ we can apply the case we have just proved, and find that for almost all x_2 the set E^{x_2} is measurable and $\mu_1(E^{x_2}) = \mu_1(F^{x_2}) - \mu_1(Z^{x_2})$. From this the proposition follows.

We now obtain the main result, generalizing Fubini's theorem in Chapter 2.

Theorem 3.3 *In the setting above, suppose $f(x_1, x_2)$ is an integrable function on $(X_1 \times X_2, \mu_1 \times \mu_2)$.*

(i) *For almost every $x_2 \in X_2$, the slice $f^{x_2}(x_1) = f(x_1, x_2)$ is integrable on (X_1, μ_1).*

(ii) *$\int_{X_1} f(x_1, x_2)\, d\mu_1$ is an integrable function on X_2.*

(iii) *$\int_{X_2} \left(\int_{X_1} f(x_1, x_2)\, d\mu_1 \right) d\mu_2 = \int_{X_1 \times X_2} f \; d\mu_1 \times \mu_2$.*

Proof. Note that if the desired conclusions hold for finitely many functions, they also hold for their linear combinations. In particular it suffices to assume that f is non-negative. When $f = \chi_E$, where E is a set of finite measure, what we wish to prove is contained in Proposition 3.2. Hence the desired result also holds for simple functions. Therefore by the monotone convergence theorem it is established for all non-negative functions, and the theorem is proved.

We remark that in general the product space $(X, \mathcal{M}, \mu)$ constructed above is not complete. However, if we define the completed space $(\overline{X}, \overline{\mathcal{M}}, \mu)$ as in Exercise 2, the theorem continues to hold in this completed space. The proof requires only a simple modification of the argument in Proposition 3.2.

3.2 Integration formula for polar coordinates

The polar coordinates of a point $x \in \mathbb{R}^d - \{0\}$ are the pair (r, γ), where $0 < r < \infty$ and γ belongs to the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d,\ |x| = 1\}$. These are determined by

$$r = |x|, \quad \gamma = \frac{x}{|x|}, \qquad \text{and reciprocally by } x = r\gamma. \tag{8}$$

Our intention here is to deal with the formula that, with appropriate definitions and under suitable hypotheses, states:

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_{S^{d-1}} \left(\int_0^\infty f(r\gamma) r^{d-1}\,dr \right) d\sigma(\gamma). \tag{9}$$

For this we consider the following pair of measure spaces. First, $(X_1, \mathcal{M}_1, \mu_1)$, where $X_1 = (0,\infty)$, $\mathcal{M}_1$ is the collection of Lebesgue measurable sets in $(0,\infty)$, and $d\mu_1(r) = r^{d-1}dr$ in the sense that $\mu_1(E) = \int_E r^{d-1}\,dr$. Next, X_2 is the unit sphere S^{d-1}, and the measure μ_2 is the one in effect determined by (9) with $\mu_2 = \sigma$. Indeed given any set $E \subset S^{d-1}$ we let $\tilde{E} = \{x \in \mathbb{R}^d : \ x/|x| \in E,\ 0 < |x| < 1\}$ be the "sector" in the unit ball whose "end-points" are in E. We shall say $E \in \mathcal{M}_2$ exactly when $\tilde{E}$ is a Lebesgue measurable subset of $\mathbb{R}^d$, and define $\mu_2(E) = \sigma(E) = d \cdot m(\tilde{E})$, where m is Lebesgue measure in $\mathbb{R}^d$.

With this it is clear that both $(X_1, \mathcal{M}_1, \mu_1)$ and $(X_2, \mathcal{M}_2, \mu_2)$ satisfy all the properties of complete and σ-finite measure spaces. We note also that the sphere S^{d-1} has a metric on it given by $d(\gamma,\gamma') = |\gamma - \gamma'|$, for $\gamma, \gamma' \in S^{d-1}$. If E is an open set (with respect to this metric) in S^{d-1}, then $\tilde{E}$ is open in $\mathbb{R}^d$, and hence E is a measurable set in S^{d-1}.

Theorem 3.4 *Suppose f is an integrable function on $\mathbb{R}^d$. Then for almost every $\gamma \in S^{d-1}$ the slice f^γ defined by $f^\gamma(r) = f(r\gamma)$ is an integrable function with respect to the measure $r^{d-1}\,dr$. Moreover, $\int_0^\infty f^\gamma(r) r^{d-1}\,dr$ is integrable on S^{d-1} and the identity* (9) *holds.*

There is a corresponding result with the order of integration of r and γ reversed.

Proof. We consider the product measure $\mu = \mu_1 \times \mu_2$ on $X_1 \times X_2$ given by Theorem 3.3. Since the space $X_1 \times X_2 = \{(r,\gamma) : \ 0 < r < \infty \text{ and } \gamma \in S^{d-1}\}$ can be identified with $\mathbb{R}^d - \{0\}$, we can think of μ as a measure of the latter space, and our main task is to identify it with the (restriction of) Lebesgue measure on that space. We claim first that

$$m(E) = \mu(E) \tag{10}$$

whenever E is a measurable rectangle $E = E_1 \times E_2$, and in this case $\mu(E) = \mu_1(E_1)\mu_2(E_2)$. In fact this holds for E_2 an arbitrary measurable subset of S^{d-1} and $E_1 = (0,1)$, because then $E = E_1 \times E_2$ is the sector $\tilde{E}_2$, while $\mu_1(E_1) = 1/d$.

Because of the relative dilation-invariance of Lebesgue measure, (10) also holds when $E = (0,b) \times E_2$, $b > 0$. A simple limiting argument then proves the result for sets $E_1 = (0,a]$, and by subtraction to all open

intervals $E_1 = (a, b)$, and thus for all open sets. Thus we have $m(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2)$ for all open sets E_1, and hence for all closed sets, and therefore for all Lebesgue measurable sets. (In fact, we can find sets $F_1 \subset E_1 \subset \mathcal{O}_1$ with F_1 closed and $\mathcal{O}_1$ open, such that $m_1(\mathcal{O}_1) - \epsilon \leq m_1(E_1) \leq m_1(F_1) + \epsilon$, and apply the above to $F_1 \times E_2$ and $\mathcal{O}_1 \times E_2$.) So we have established the identity (10) for all measurable rectangles and as a result for all finite unions of measurable rectangles. This is the algebra $\mathcal{A}$ that occurs in the proof of Theorem 3.3, and hence by the uniqueness in Theorem 1.5, the identity extends to the σ-algebra generated by $\mathcal{A}$, which is the σ-algebra $\mathcal{M}$ on which the measure μ is defined. To summarize, whenever $E \in \mathcal{M}$, the assertion (9) holds for $f = \chi_E$.

To go further we note that any open set in $\mathbb{R}^d - \{0\}$ can be written as a countable union of rectangles $\bigcup_{j=1}^{\infty} A_j \times B_j$, where A_j and B_j are open in $(0, \infty)$ and S^{d-1}, respectively. (This small technical point is taken up in Exercise 12.) It follows that any open set is in $\mathcal{M}$, and therefore so is any Borel set. Thus (9) is valid for χ_E whenever E is any Borel set in $\mathbb{R}^d - \{0\}$. The result then goes over to any Lebesgue set $E' \subset \mathbb{R}^d - \{0\}$, since such a set can be written as a disjoint union $E' = E \cup Z$, where E is a Borel set and $Z \subset F$, with F a Borel set of measure zero. To finish the proof we follow the familiar steps of deducing (9) for simple functions, and then by monotonic convergence for non-negative integrable functions, and from that for the general case.

3.3 Borel measures on $\mathbb{R}$ and the Lebesgue-Stieltjes integral

The Stieltjes integral was introduced to provide a generalization of the Riemann integral $\int_a^b f(x)\,dx$, where the increments dx were replaced by the increments $dF(x)$ for a given increasing function F on $[a, b]$. We wish to pursue this idea from the general point of view taken in this chapter. The question that is then raised is that of characterizing the measures on $\mathbb{R}$ that arise in this way, and in particular measures defined on the Borel sets on the real line.

To have a unique correspondence between measures and increasing functions as we shall have below, we need first to normalize these functions appropriately. Recall that an increasing function F can have at most a countable number of discontinuities. If x_0 is such a discontinuity, then

$$\lim_{\substack{x < x_0 \\ x \to x_0}} F(x) = F(x_0^-) \quad \text{and} \quad \lim_{\substack{x > x_0 \\ x \to x_0}} F(x) = F(x_0^+)$$

both exist, while $F(x_0^-) < F(x_0^+)$ and $F(x_0)$ is some value between $F(x_0^-)$ and $F(x_0^+)$. We shall now modify F at x_0, if necessary, by setting $F(x_0) = F(x_0^+)$, and we do this for every point of discontinuity. The function F so obtained is now still increasing, yet right-continuous at every point, and we say such functions are **normalized**. The main result is then as follows.

Theorem 3.5 *Let F be an increasing function on $\mathbb{R}$ that is normalized. Then there is a unique measure μ (also denoted by dF) on the Borel sets $\mathcal{B}$ on $\mathbb{R}$ such that $\mu((a,b]) = F(b) - F(a)$ if $a < b$. Conversely, if μ is a measure on $\mathcal{B}$ that is finite on bounded intervals, then F defined by $F(x) = \mu((0,x])$, $x > 0$, $F(0) = 0$ and, $F(x) = -\mu((-x,0])$, $x < 0$, is increasing and normalized.*

Before we come to the proof, we remark that the condition that μ be finite on bounded intervals is crucial. In fact, the Hausdorff measures that will be considered in the next chapter provide examples of Borel measures on $\mathbb{R}$ of a very different character from those treated in the theorem.

Proof. We define a function μ_* on all subsets of $\mathbb{R}$ by

$$\mu_*(E) = \inf \sum_{j=1}^{\infty} (F(b_j) - F(a_j)),$$

where the infimum is taken over all coverings of E of the form $\bigcup_{j=1}^{\infty}(a_j, b_j]$.

It is easy to verify that μ_* is an exterior measure on $\mathbb{R}$. We observe next that $\mu_*((a,b]) = (F(b) - F(a))$, if $a < b$. Clearly $\mu_*((a,b]) \leq F(b) - F(a)$, since $(a,b]$, then covers itself. Next, suppose that $\bigcup_{j=1}^{\infty}(a_j, b_j]$ covers $(a,b]$; then it covers $[a', b]$ for any $a < a' < b$. However, by the right-continuity of F, if $\epsilon > 0$ is given, we can always choose $b_j' > b_j$ such that $F(b_j') \leq F(b_j) + \epsilon/2^j$. Now the union of open intervals $\bigcup_{j=1}^{\infty}(a_j, b_j')$ covers $[a', b]$. By the compactness of this interval, $\bigcup_{j=1}^{N}(a_j, b_j')$ covers $[a', b]$ for some N. Thus since F is increasing we have

$$\begin{aligned} F(b) - F(a') \leq \sum_{j=1}^{N} F(b_j') - F(a_j) &\leq \sum_{j=1}^{N} (F(b_j) - F(a_j) + \epsilon/2^j) \\ &\leq \mu_*((a,b]) + \epsilon. \end{aligned}$$

Thus letting $a' \to a$, and using the right-continuity of F again, we see that $F(b) - F(a) \leq \mu_*((a,b]) + \epsilon$. Since ϵ was arbitrary this then proves $F(b) - F(a) = \mu_*((a,b])$.

Next we show that μ_* is a metric exterior measure (for the usual metric $d(x, x') = |x - x'|$ on the real line). Since μ_* is an exterior measure we have $\mu_*(E_1 \cup E_2) \le \mu_*(E_1) + \mu_*(E_2)$; thus it suffices to see that the reverse inequality holds whenever $d(E_1, E_2) \ge \delta$, for some $\delta > 0$.

Suppose that we are given a positive ϵ, and that $\bigcup_{j=1}^{\infty}(a_j, b_j]$ is a covering of $E_1 \cup E_2$ such that

$$\sum_{j=1}^{\infty} F(b_j) - F(a_j) \le \mu_*(E_1 \cup E_2) + \epsilon.$$

We may assume, after subdividing the intervals $(a_j, b_j]$ into smaller half-open intervals, that each interval in the covering has length less than δ. When this is so each interval can intersect at most one of the two sets E_1 or E_2. If we denote by J_1 and J_2 the sets of those indices for which $(a_j, b_j]$ intersects E_1 and E_2, respectively, then $J_1 \cap J_2$ is empty; moreover, we have $E_1 \subset \bigcup_{j \in J_1}(a_j, b_j]$ as well as $E_2 \subset \bigcup_{j \in J_2}(a_j, b_j]$. Therefore

$$\begin{aligned} \mu_*(E_1) + \mu_*(E_2) &\le \sum_{j \in J_1} F(b_j) - F(a_j) + \sum_{j \in J_2} F(b_j) - F(a_j) \\ &\le \sum_{j=1}^{\infty} F(b_j) - F(a_j) \le \mu_*(E_1 \cup E_2) + \epsilon. \end{aligned}$$

Since ϵ was arbitrary, we see that $\mu_*(E_1) + \mu_*(E_2) \le \mu_*(E_1 \cup E_2)$, as we intended to show.

We can now invoke Theorem 1.5. This guarantees the existence of a measure μ for which the Borel sets are measurable; moreover, we have $\mu((a, b]) = F(b) - F(a)$, since clearly $(a, b])$ is a Borel set and we have previously seen that $\mu_*((a, b]) = F(b) - F(a)$.

To prove that μ is the unique Borel measure on $\mathbb{R}$ for which $\mu((a, b]) = F(b) - F(a)$, let us suppose that ν is another Borel measure with this property. It now suffices to show that $\nu = \mu$ on all Borel sets.

We can write any open interval as a disjoint union $(a, b) = \bigcup_{j=1}^{\infty}(a_j, b_j]$, by choosing $\{b_j\}_{j=1}^{\infty}$ to be a strictly increasing sequence with $a < b_j < b$, $b_j \to b$ as $j \to \infty$, and taking $a_1 = a$, $a_{j+1} = b_j$. Since ν and μ agree on each $(a_j, b_j]$, it follows that ν and μ agree on (a, b), and hence on all open intervals, and therefore on all open sets. Moreover, clearly ν and μ are finite on all bounded intervals; thus the regularity in Proposition 1.3 allows one to conclude that $\mu = \nu$ on all Borel sets.

Conversely, if we start with a Borel measure μ on $\mathbb{R}$ that is finite on bounded intervals, we can define the function F as in the statement of the theorem. Then clearly F is increasing. To see that it is right-continuous,

note that if, for instance, $x_0 > 0$, the sets $E_n = (0, x_0 + 1/n]$ decrease to $E = (0, x_0]$ as $n \to \infty$, hence $\mu(E_n) \to \mu(E)$, since $\mu(E_1) < \infty$. This means that $F(x_0 + 1/n) \to F(x_0)$. Since F is increasing, this implies that F is right-continuous at x_0. The argument for any $x_0 \leq 0$ is similar, and thus the theorem is proved.

Remarks. Several comments about the theorem are in order.

(i) Two increasing functions F and G give the same measure if $F - G$ is constant. The converse if also true because $F(b) - F(a) = G(b) - G(a)$ for all $a < b$ exactly when $F - G$ is constant.

(ii) The measure μ constructed in the proof of the theorem is defined on a larger σ-algebra than the Borel sets, and is actually complete. However, in applications, its restriction to the Borel sets often suffices.

(iii) If F is an increasing normalized function given on a closed interval $[a, b]$, we can extend it to $\mathbb{R}$ by setting $F(x) = F(a)$ for $x < a$, and $F(x) = F(b)$ for $x > b$. For the resulting measure μ, the intervals $(-\infty, a]$ and (b, ∞) have measure zero. One then often writes

$$\int_{\mathbb{R}} f(x)\, d\mu(x) = \int_a^b f(x)\, dF(x),$$

for every f that is integrable with respect to μ. If F arises from an increasing function F_0 defined on $\mathbb{R}$, one may wish to account for the possible jump of F_0 at a. In this case it is sometimes useful to define

$$\int_{a^-}^b f(x)\, dF(x) \quad \text{as} \quad \int_a^b f(x)\, d\mu_0(x),$$

where μ_0 is the measure on $\mathbb{R}$ corresponding to F_0.

(iv) Note that the above definition of the Lebesgue-Stieltjes integral extends to the case when F is of bounded variation. Indeed suppose F is a complex-valued function on $[a, b]$ such that $F = \sum_{j=1}^4 \epsilon_j F_j$, where each F_j is increasing and normalized, and ϵ_j are ± 1 or $\pm i$. Then we can define $\int_a^b f(x)\, dF(x)$ as $\sum_{j=1}^4 \epsilon_j \int_a^b f(x)\, dF_j(x)$; here we require that f be integrable with respect to the Borel measure $\mu = \sum_{j=1}^4 \mu_j$, where μ_j is the measure corresponding to F_j.

(v) The value of these integrals can be calculated more directly in the following cases.

(a) If F is an absolutely continuous function on $[a, b]$, then

$$\int_a^b f(x)\, dF(x) = \int_a^b f(x) F'(x)\, dx$$

for every Borel measurable function f that is integrable with respect to $\mu = dF$.

(b) Suppose F is a pure jump function as in Section 3.3, Chapter 3, with jumps $\{\alpha_n\}_{n=1}^\infty$ at the points $\{x_n\}_{n=1}^\infty$. Then whenever f is, say, continuous and vanishes outside some finite interval we have

$$\int_a^b f(x)\, dF(x) = \sum_{n=1}^\infty f(x_n)\alpha_n.$$

In particular, for the measure μ we have $\mu(\{x_n\}) = \alpha_n$ and $\mu(E) = 0$ for all sets that do not contain any of the x_n.

(c) A special instance arises when $F = H$, the Heaviside function defined by $H(x) = 1$ for $x \geq 0$, and $H(x) = 0$ for $x < 0$. Then

$$\int_{-\infty}^\infty f(x)\, dH(x) = f(0),$$

which is another expression for the Dirac delta function arising in Section 2 of Chapter 3.

Further details about (v) can be found in Exercise 11.

4 Absolute continuity of measures

The generalization of the notion of absolute continuity considered in Chapter 3 requires that we extend the ideas of a measure to encompass set functions that may be positive or negative. We describe this notion first.

4.1 Signed measures

Loosely speaking, a signed measure possesses all the properties of a measure, except that it may take positive or negative values. More precisely, a **signed measure** ν on a σ-algebra $\mathcal{M}$ is a mapping that satisfies:

(i) The set function ν is extended-valued in the sense that $-\infty < \nu(E) \leq \infty$ for all $E \in \mathcal{M}$.

(ii) If $\{E_j\}_{j=1}^{\infty}$ are disjoint subsets of $\mathcal{M}$, then

$$\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j).$$

Note that for this to hold the sum $\sum \nu(E_j)$ must be independent of the rearrangements of terms, so that if $\nu(\bigcup_{j=1}^{\infty} E_j)$ is finite, it implies that the sum converges absolutely.

Examples of signed measures arise naturally if we drop the assumption that f be non-negative in the expression

$$\nu(E) = \int_E f \, d\mu,$$

where $(X, \mathcal{M}, \mu)$ is a measure space and f is μ-measurable. In fact, to ensure that ν satisfies (i) and (ii) the function f is required to be "integrable" with respect to μ in the extended sense that $\int f^- \, d\mu$ must be finite, while $\int f^+ \, d\mu$ may be infinite.

Given a signed measure ν on $(X, \mathcal{M})$ it is always possible to find a (positive) measure μ that dominates ν, in the sense that

$$\nu(E) \leq \mu(E) \quad \text{for all } E,$$

and that in addition is the "smallest" μ that has this property.

The construction is in effect an abstract version of the decomposition of a function of bounded variation as the difference of two increasing functions, as carried out in Chapter 3. We proceed as follows. We define a function $|\nu|$ on $\mathcal{M}$, called the **total variation** of ν, by

$$|\nu|(E) = \sup \sum_{j=1}^{\infty} |\nu(E_j)|,$$

where the supremum is taken over all partitions of E, that is, over all countable unions $E = \bigcup_{j=1}^{\infty} E_j$, where the sets E_j are disjoint and belong to $\mathcal{M}$.

The fact that $|\nu|$ is actually additive is not obvious, and is given in the proof below.

Proposition 4.1 *The total variation $|\nu|$ of a signed measure ν is itself a (positive) measure that satisfies $\nu \leq |\nu|$.*

Proof. Suppose $\{E_j\}_{j=1}^{\infty}$ is a countable collection of disjoints sets in $\mathcal{M}$, and let $E = \bigcup E_j$. It suffices to prove:

$$\sum |\nu|(E_j) \leq |\nu|(E) \quad \text{and} \quad |\nu|(E) \leq \sum |\nu|(E_j). \tag{11}$$

Let α_j be a real number that satisfies $\alpha_j < |\nu|(E_j)$. By definition, each E_j can be written as $E_j = \bigcup_i F_{i,j}$, where the $F_{i,j}$ are disjoint, belong to $\mathcal{M}$, and

$$\alpha_j \leq \sum_{i=1}^{\infty} |\nu(F_{i,j})|.$$

Since $E = \bigcup_{i,j} F_{i,j}$, we have

$$\sum \alpha_j \leq \sum_{j,i} |\nu(F_{i,j})| \leq |\nu|(E).$$

Consequently, taking the supremum over the numbers α_j gives the first inequality in (11).

For the reverse inequality, let F_k be any other partition of E. For a fixed k, $\{F_k \cap E_j\}_j$ is a partition of F_k, so

$$\sum_k |\nu(F_k)| = \sum_k \left| \sum_j \nu(F_k \cap E_j) \right|,$$

since ν is a signed measure. An application of the triangle inequality and the fact that $\{F_k \cap E_j\}_k$ is a partition of E_j gives

$$\begin{aligned}
\sum_k |\nu(F_k)| &\leq \sum_k \sum_j |\nu(F_k \cap E_j)| \\
&= \sum_j \sum_k |\nu(F_k \cap E_j)| \\
&\leq \sum_j |\nu|(E_j).
\end{aligned}$$

Since $\{F_k\}$ was an arbitrary partition of E, we obtain the second inequality in (11) and the proof is complete.

It is now possible to write ν as the difference of two (positive) measures. To see this, we define the **positive variation** and **negative variation** of ν by

$$\nu^+ = \frac{1}{2}(|\nu| + \nu) \quad \text{and} \quad \nu^- = \frac{1}{2}(|\nu| - \nu).$$

By the proposition we see that ν^+ and ν^- are measures, and they clearly satisfy

$$\nu = \nu^+ - \nu^- \quad \text{and} \quad |\nu| = \nu^+ + \nu^-.$$

In the above if $\nu(E) = \infty$ for a set E, then $|\nu|(E) = \infty$, and $\nu^-(E)$ is defined to be zero.

We also make the following definition: we say that the signed measure ν is **σ-finite** if the measure $|\nu|$ is σ-finite. Since $\nu \leq |\nu|$ and $|-\nu| = |\nu|$, we find that

$$-|\nu| \leq \nu \leq |\nu|.$$

As a result, if ν is σ-finite, then so are ν^+ and ν^-.

4.2 Absolute continuity

Given two measures defined on a common σ-algebra we describe here the relationships that can exist between them. More concretely, consider two measures ν and μ defined on the σ-algebra $\mathcal{M}$; two extreme scenarios are

(a) ν and μ are "supported" on separate parts of $\mathcal{M}$.

(b) The support of ν is an essential part of the support of μ.

Here we adopt the terminology that the measure ν is **supported** on a set A, if $\nu(E) = \nu(E \cap A)$ for all $E \in \mathcal{M}$.

The Lebesgue-Radon-Nikodym theorem below states that in a precise sense the relationship between any two measures ν and μ is a combination of the above two possibilities.

Mutually singular and absolutely continuous measures

Two signed measures ν and μ on $(X, \mathcal{M})$ are **mutually singular** if there are disjoint subsets A and B in $\mathcal{M}$ so that

$$\nu(E) = \nu(A \cap E) \quad \text{and} \quad \mu(E) = \mu(B \cap E) \quad \text{for all } E \in \mathcal{M}.$$

Thus ν and μ are supported on disjoint subsets. We use the symbol $\nu \perp \mu$ to denote the fact that the measures are mutually singular.

In contrast, if ν is a signed measure and μ a (positive) measure on $\mathcal{M}$, we say that ν is **absolutely continuous** with respect to μ if

$$\nu(E) = 0 \quad \text{whenever } E \in \mathcal{M} \text{ and } \mu(E) = 0. \tag{12}$$

Thus if ν is supported in a set A, then A must be an essential part of the support of μ in the sense that $\mu(A) > 0$. We use the symbol $\nu \ll \mu$ to indicate that ν is absolutely continuous with respect to μ. Note that if ν and μ are mutually singular, and ν is also absolutely continuous with respect to μ, then ν vanishes identically.

An important example is given by integration with respect to μ. Indeed, if $f \in L^1(X, \mu)$, or if f is merely integrable in the extended sense (where $\int f^- < \infty$, but possibly $\int f^+ = \infty$), then the signed measure ν defined by

$$\nu(E) = \int_E f \, d\mu \tag{13}$$

is absolutely continuous with respect to μ. We shall use the shorthand $d\nu = f d\mu$ to indicate that ν is defined by (13).

This is a variant of the notion of absolute continuity that arose in Chapter 3 in the special case of $\mathbb{R}$ (with $\mathcal{M}$ the Lebesgue measurable sets and $d\mu = dx$ the Lebesgue measure). In fact, with ν defined by (13) and f an integrable function, we saw that in place of (12) we had the following stronger assertion:

(14)

For each $\epsilon > 0$*, there is a* $\delta > 0$ *such that* $\mu(E) < \delta$ *implies* $|\nu(E)| < \epsilon$.

In the general situation the relation between the two conditions (12) and (14) is clarified by the following observation.

Proposition 4.2 *The assertion* (14) *implies* (12). *Conversely, if* $|\nu|$ *is a finite measure, then* (12) *implies* (14).

That (12) is a consequence of (14) is obvious because $\mu(E) = 0$ gives $|\nu(E)| < \epsilon$ for every $\epsilon > 0$. To prove the converse, it suffices to consider the case when ν is positive, upon replacing ν by $|\nu|$. We then assume that (14) does not hold. This means that it fails for some fixed $\epsilon > 0$. Hence for each n, there is a measurable set E_n with $\mu(E_n) < 2^{-n}$ while $\nu(E_n) \geq \epsilon$. Now let $E^* = \limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n^*$, where $E_n^* = \bigcup_{k \geq n} E_k$. Then since $\mu(E_n^*) \leq \sum_{k\geq n} 1/2^k = 1/2^{n-1}$, and the decreasing sets $\{E_k^*\}$ are contained in a set of finite measure (E_1^*), we get $\mu(E^*) = 0$. However $\nu(E_n^*) \geq \nu(E_n) \geq \epsilon$, and the ν measure is assumed finite. So $\nu(E^*) = \lim_{n\to\infty} \nu(E_n^*) \geq \epsilon$, which gives a contradiction.

After these preliminaries we can come to the main result. It guarantees among other things a converse to the representation (13); it was proved in the case of $\mathbb{R}$ by Lebesgue, and in the general case by Radon and Nikodym.

Theorem 4.3 *Suppose μ is a σ-finite positive measure on the measure space $(X, \mathcal{M})$ and ν a σ-finite signed measure on $\mathcal{M}$. Then there exist unique signed measures ν_a and ν_s on $\mathcal{M}$ such that $\nu_a \ll \mu$, $\nu_s \perp \mu$ and $\nu = \nu_a + \nu_s$. In addition, the measure ν_a takes the form $d\nu_a = f d\mu$; that is,*

$$\nu_a(E) = \int_E f(x)\, d\mu(x)$$

for some extended μ-integrable function f.

Note the following consequence. If ν is absolutely continuous with respect to μ, then $d\nu = f d\mu$, and this assertion can be viewed as a generalization of Theorem 3.11 in Chapter 3.

There are several known proofs of the above theorem. The argument given below, due to von Neumann, has the virtue that it exploits elegantly the application of a simple Hilbert space idea.

We start with the case when both ν and μ are positive and finite. Let $\rho = \nu + \mu$, and consider the transformation on $L^2(X, \rho)$ defined by

$$\ell(\psi) = \int_X \psi(x)\, d\nu(x).$$

The mapping ℓ defines a bounded linear functional on $L^2(X, \rho)$ since

$$\begin{aligned} |\ell(\psi)| \leq \int_X |\psi(x)|\, d\nu(x) &\leq \int_X |\psi(x)|\, d\rho(x) \\ &\leq (\rho(X))^{1/2} \left(\int_X |\psi(x)|^2\, d\rho(x) \right)^{1/2}, \end{aligned}$$

where the last inequality follows by the Cauchy-Schwarz inequality. But $L^2(X, \rho)$ is a Hilbert space, so the Riesz representation theorem (in Chapter 4) guarantees the existence of $g \in L^2(X, \rho)$ such that

$$\int_X \psi(x)\, d\nu(x) = \int_X \psi(x) g(x)\, d\rho(x) \quad \text{for all } \psi \in L^2(X, \rho). \tag{15}$$

If $E \in \mathcal{M}$ with $\rho(E) > 0$, when we set $\psi = \chi_E$ in (15) and recall that $\nu \leq \rho$, we find

$$0 \leq \frac{1}{\rho(E)} \int_E g(x)\, d\rho(x) \leq 1,$$

from which we conclude that $0 \leq g(x) \leq 1$ for a.e. x (with respect to the measure ρ). In fact, $0 \leq \int_E g(x)\, d\rho(x)$ for all sets $E \in \mathcal{M}$ implies that

$g(x) \geq 0$ almost everywhere. In the same way, $0 \leq \int_E (1 - g(x))\, d\rho(x)$ for all $E \in \mathcal{M}$ guarantees that $g(x) \leq 1$ almost everywhere. Therefore we may clearly assume $0 \leq g(x) \leq 1$ for all x without disturbing the identity (15), which we rewrite as

$$\int \psi(1 - g)\, d\nu = \int \psi g\, d\mu. \tag{16}$$

Consider now the two sets

$$A = \{x \in X :\ 0 \leq g(x) < 1\} \quad \text{and} \quad B = \{x \in X :\ g(x) = 1\},$$

and define two measures ν_a and ν_s on $\mathcal{M}$ by

$$\nu_a(E) = \nu(A \cap E) \quad \text{and} \quad \nu_s(E) = \nu(B \cap E).$$

To see why $\nu_s \perp \mu$, it suffices to note that setting $\psi = \chi_B$ in (16) gives

$$0 = \int \chi_B\, d\mu = \mu(B).$$

Finally, we set $\psi = \chi_E(1 + g + \cdots + g^n)$ in (16) :

$$\int_E (1 - g^{n+1})\, d\nu = \int_E g(1 + \cdots + g^n)\, d\mu. \tag{17}$$

Since $(1 - g^{n+1})(x) = 0$ if $x \in B$, and $(1 - g^{n+1})(x) \to 1$ if $x \in A$, the dominated convergence theorem implies that the left-hand side of (17) converges to $\nu(A \cap E) = \nu_a(E)$. Also, $1 + g + \cdots + g^n$ converges to $\frac{1}{1-g}$, so we find in the limit that

$$\nu_a(E) = \int_E f\, d\mu, \qquad \text{where } f = \tfrac{g}{1-g}.$$

Note that $f \in L^1(X, \mu)$, since $\nu_a(X) \leq \nu(X) < \infty$. If μ and ν are σ-finite and positive we may clearly find sets $E_j \in \mathcal{M}$ such that $X = \bigcup E_j$ and

$$\mu(E_j) < \infty, \qquad \nu(E_j) < \infty \qquad \text{for all } j.$$

We may define positive and finite measures on $\mathcal{M}$ by

$$\mu_j(E) = \mu(E \cap E_j) \quad \text{and} \quad \nu_j(E) = \nu(E \cap E_j),$$

and then we can write for each j, $\nu_j = \nu_{j,a} + \nu_{j,s}$ where $\nu_{j,s} \perp \mu_j$ and $\nu_{j,a} = f_j\, d\mu_j$. Then it suffices to set

$$f = \sum f_j, \qquad \nu_s = \sum \nu_{j,s}, \qquad \text{and} \qquad \nu_a = \sum \nu_{j,a}.$$

Finally, if ν is signed, then we apply the argument separately to the positive and negative variations of ν.

To prove the uniqueness of the decomposition, suppose we also have $\nu = \nu_a' + \nu_s'$, where $\nu_a' \ll \mu$ and $\nu_s' \perp \mu$. Then

$$\nu_a - \nu_a' = \nu_s' - \nu_s.$$

The left-hand side is absolutely continuous with respect to μ, and the right-hand side is singular with respect to μ. Thus both sides are zero and the theorem is proved.

5* Ergodic theorems

Ergodic theory had its beginnings in certain problems in statistical mechanics studied in the late nineteenth century. Since then it has grown rapidly and has gained wide influence in a number of mathematical disciplines, in particular those related to dynamical systems and probability theory. It is not our purpose to try to give an account of this broad and fascinating theory. Rather, we restrict our presentation to some of the basic limit theorems that lie at its foundation. These theorems are most naturally formulated in the general context of abstract measure spaces, and thus for us they serve as excellent illustrations of the general framework developed in this chapter.

The setting for the theory is a σ-finite measure space $(X, \mathcal{M}, \mu)$ endowed with a mapping $\tau : X \to X$ such that whenever E is a measurable subset of X, then so is $\tau^{-1}(E)$, and $\mu(\tau^{-1}(E)) = \mu(E)$. Here $\tau^{-1}(E)$ is the pre-image of E under τ; that is, $\tau^{-1}(E) = \{x \in X : \tau(x) \in E\}$. A mapping τ with these properties is called a **measure-preserving transformation**. If in addition for such a τ we have the feature that it is a bijection and τ^{-1} is also a measure-preserving transformation, then τ is referred to as a **measure-preserving isomorphism**.

Let us note that if τ is a measure-preserving transformation, then $f(\tau(x))$ is measurable if $f(x)$ is measurable, and is integrable if f is integrable; moreover, then

$$\int_X f(\tau(x))\, d\mu(x) = \int_X f(x)\, d\mu(x). \tag{18}$$

Indeed, if χ_E is the characteristic function of the set E, we note that $\chi_E(\tau(x)) = \chi_{\tau^{-1}(E)}(x)$, and so the assertion holds for characteristic functions of measurable sets and thus for simple functions, and hence by the usual limiting arguments for all non-negative measurable functions, and

then integrable functions. For later purposes we record here an equivalent statement: whenever f is a real-valued measurable function and α is any real number, then

$$\mu(\{x : f(x) > \alpha\}) = \mu(\{x : f(\tau(x)) > \alpha\}).$$

Before we proceed further, we describe several examples of measure-preserving transformations:

(i) Here $X = \mathbb{Z}$, the integers, with μ its counting measure; that is, $\mu(E) = \#(E) =$ the number of integers in E, for any $E \subset \mathbb{Z}$. We define τ to be the unit translation, $\tau : n \mapsto n + 1$. Note that τ gives a measure-preserving isomorphism of $\mathbb{Z}$.

(ii) Another easy example is $X = \mathbb{R}^d$ with Lebesgue measure, and τ a translation, $\tau : x \mapsto x + h$ for some fixed $h \in \mathbb{R}^d$. This is of course a measure-preserving isomorphism. (See the section on invariance properties of the Lebesgue measure in Chapter 1.)

(iii) Here X is the unit circle, given as $\mathbb{R}/\mathbb{Z}$, with the measure induced from Lebesgue measure on $\mathbb{R}$. That is, we may realize X as the unit interval $(0, 1]$, and take μ to be the Lebesgue measure restricted to this interval. For any real number α, the translation $x \mapsto x + \alpha$, taken modulo $\mathbb{Z}$, is well defined on $X = \mathbb{R}/\mathbb{Z}$, and is measure-preserving. (See the related Exercise 3 in Chapter 2.) It can be interpreted as a rotation of the circle by angle $2\pi\alpha$.

(iv) In this example X is again $(0, 1]$ with Lebesgue measure μ, but τ is the doubling map $\tau(x) = 2x \mod 1$. It is easy to verify that τ is a measure-preserving transformation. Indeed, any set $E \subset (0, 1]$ has two pre-images E_1 and E_2, the first in $(0, 1/2]$ and the second in $(1/2, 1]$, both of measure $\mu(E)/2$, if E is measurable. (See Figure 1.) However, τ is not an isomorphism, since τ is not injective.

(v) A trickier example is given by the transformation that is key in the theory of continued fractions. Here $X = [0, 1)$ and τ is defined by $\tau(x) = \langle 1/x \rangle$, the fractional part of $1/x$; when $x = 0$ we set $\tau(0) = 0$. Gauss observed, in effect, that the measure $d\mu = \frac{1}{1+x}\, dx$ is preserved by the transformation τ. Note that each $x \in (0, 1)$ has infinitely many pre-images under τ; that is, the sequence $\{1/(x + k)\}_{k=1}^{\infty}$. More about this example can be found in Problems 8 through 10 below.

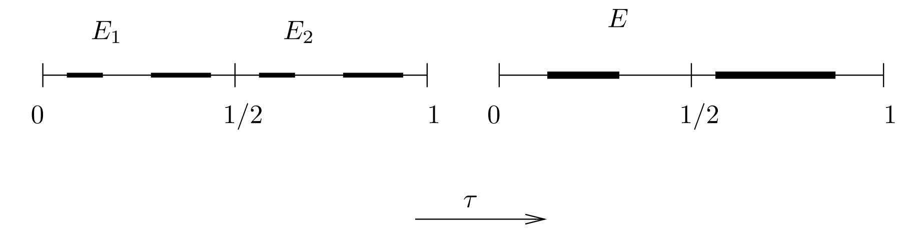


Figure 1. Pre-images E_1 and E_2 under the doubling map

Having pointed out these examples, we can now return to the general theory. The notions described above are of interest, in part, because they abstract the idea of a dynamical system, one whose totality of states is represented by the space X, with each point $x \in X$ giving a particular state of the system. The mapping $\tau : X \to X$ then describes the transformation of the system after a unit of time has elapsed. For such a system there is often associated a notion of "volume" or "mass" that is unchanged by the evolution, and this is the role of the invariant measure μ. The iterates, $\tau^n = \tau \circ \tau \circ \cdots \circ \tau$ (n times) describe the evolution of the system after n units of time, and a principal concern is the average behaviour, as $n \to \infty$, of various quantities associated with the system. Thus one is led to study averages

$$A_n(f)(x) = \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x)), \tag{19}$$

and their limits as $n \to \infty$. To this we now turn.

5.1 Mean ergodic theorem

The first theorem dealing with the averages (19) that we consider is purely Hilbert-space in character. Historically it preceded both Theorems 5.3 and 5.4 which will be proved below.

For the specific application of the theorem below, one takes the Hilbert space $\mathcal{H}$ to be $L^2(X, \mathcal{M}, \mu)$. Given the measure-preserving transformation τ on X, we define the linear operator T on $\mathcal{H}$ by

$$T(f)(x) = f(\tau(x)). \tag{20}$$

Then T is an isometry; that is,

$$\|Tf\| = \|f\|, \tag{21}$$

where $\|\cdot\|$ denotes the Hilbert space (that is, the L^2) norm. This is clear from (18) with f replaced by $|f|^2$. Observe that if τ were also supposed to be a measure-preserving isomorphism, then T would be invertible and hence unitary; but we do not assume this.

Now with T as above, consider the subspace S of **invariant vectors**, $S = \{f \in \mathcal{H} : \ T(f) = f\}$. Clearly, because of (21), the subspace S is closed. Let P denote the orthogonal projection on this subspace. The theorem that follows deals with the "mean" convergence, meaning convergence in the norm.

Theorem 5.1 *Suppose T is an isometry of the Hilbert space $\mathcal{H}$, and let P be the orthogonal projection on the subspace of the invariant vectors of T. Let $A_n = \frac{1}{n}(I + T + T^2 + \cdots + T^{n-1})$. Then for each $f \in \mathcal{H}$, $A_n(f)$ converges to $P(f)$ in norm, as $n \to \infty$.*

Together with the subspace S defined above we consider the subspaces $S_* = \{f \in \mathcal{H} : \ T^*(f) = f\}$ and $S_1 = \{f \in \mathcal{H} : \ f = g - Tg, \ g \in \mathcal{H}\}$; here T^* denotes the adjoint of T. Then S_*, like S, is closed, but S_1 is not necessarily closed. We denote its closure by $\overline{S_1}$. The proof of the theorem is based on the following lemma.

Lemma 5.2 *The following relations hold among the subspaces S, S_*, and $\overline{S_1}$.*

(i) $S = S_*$.

(ii) *The orthogonal complement of $\overline{S_1}$ is S.*

Proof. First, since T is an isometry, we have that $(Tf, Tg) = (f, g)$ for all $f, g \in \mathcal{H}$, and thus $T^*T = I$. (See Exercise 22 in Chapter 4.) So if $Tf = f$ then $T^*Tf = T^*f$, which means that $f = T^*f$. To prove the converse inclusion, assume $T^*f = f$. As a consequence $(f, T^*f - f) = 0$, and thus $(f, T^*f) - (f, f) = 0$; that is, $(Tf, f) = \|f\|^2$. However, $\|Tf\| = \|f\|$, so we have in the above an instance of equality for the Cauchy-Schwarz inequality. As a result of Exercise 2 in Chapter 4 we get $Tf = cf$, which by the above gives $Tf = f$. Thus part (i) is proved.

Next we observe that f is in the orthogonal complement of $\overline{S_1}$ exactly when $(f, g - Tg) = 0$, for all $g \in \mathcal{H}$. However, this means that $(f - T^*f, g) = 0$ for all g, and hence $f = T^*f$, which by part (i) means $f \in S$.

Having established the lemma we can finish the proof of the theorem. Given any $f \in \mathcal{H}$, we write $f = f_0 + f_1$, where $f_0 \in S$ and $f_1 \in \overline{S_1}$ (since S and $\overline{S_1}$ are orthogonal complements). We also fix $\epsilon > 0$ and pick $f_1' \in$

S_1 such that $\|f_1 - f_1'\| < \epsilon$. We then write

$$A_n(f) = A_n(f_0) + A_n(f_1') + A_n(f_1 - f_1'), \tag{22}$$

and consider each term separately.

For the first term, we recall that P is the orthogonal projection on S, so $P(f) = f_0$, and since $Tf_0 = f_0$ we deduce

$$A_n(f_0) = \frac{1}{n}\sum_{k=0}^{n-1} T^k(f_0) = f_0 = P(f) \qquad \text{for every } n \geq 1.$$

For the second term, we recall the definition of S_1 and pick a $g \in \mathcal{H}$ with $f_1' = g - Tg$. Thus

$$\begin{aligned} A_n(f_1') = \frac{1}{n}\sum_{k=0}^{n-1} T^k(1-T)(g) &= \frac{1}{n}\sum_{k=0}^{n-1} T^k(g) - T^{k+1}(g) \\ &= \frac{1}{n}(g - T^n(g)). \end{aligned}$$

Since T is an isometry, the above identity shows that $A_n(f_1')$ converges to 0 in the norm as $n \to \infty$.

For the last term, we use once again the fact that each T^k is an isometry to obtain

$$\|A_n(f_1 - f_1')\| \leq \frac{1}{n}\sum_{k=0}^{n-1} \|T^k(f_1 - f_1')\| \leq \|f_1 - f_1'\| < \epsilon.$$

Finally, from (22) and the above three observations, we deduce that $\limsup_{n\to\infty} \|A_n(f) - P(f)\| \leq \epsilon$, and this concludes the proof of the theorem.

5.2 Maximal ergodic theorem

We now turn to the question of almost everywhere convergence of the averages (19). As in the case of the averages that occur in the differentiation theorems of Chapter 3, the key to dealing with such pointwise limits lies in estimates for their corresponding maximal functions. In the present case this function is defined by

$$f^*(x) = \sup_{1\leq m<\infty} \frac{1}{m}\sum_{k=0}^{m-1} |f(\tau^k(x))|. \tag{23}$$

Theorem 5.3 *Whenever $f \in L^1(X, \mu)$, the maximal function $f^*(x)$ is finite for almost every x. Moreover, there is a universal constant A so that*

$$\mu(\{x : \ f^*(x) > \alpha\}) \le \frac{A}{\alpha} \, \|f\|_{L^1(X,\mu)} \quad \text{for all } \alpha > 0. \tag{24}$$

There are several proofs of this theorem. The one we choose emphasizes the close connection to the maximal function given in Section 1.1 of Chapter 3, and we shall in fact deduce the present theorem from the one-dimensional case of that chapter. This argument gives the value $A = 6$ for the constant in (24). By a different argument one can obtain $A = 1$, but this improvement is not relevant in what follows.

Before beginning the proof, we make some preliminary remarks. Note that in the present case the function f^* is automatically measurable, since it is the supremum of a countable number of measurable functions. Also, we may assume that our function f is non-negative, since otherwise we may replace it by $|f|$.

Step 1. The case when $X = \mathbb{Z}$ and $\tau : n \mapsto n + 1$.

For each function f on $\mathbb{Z}$, we consider its extension $\tilde{f}$ to $\mathbb{R}$ defined by $\tilde{f}(x) = f(n)$ for $n \le x < n + 1$, $n \in \mathbb{Z}$. (See Figure 2.)

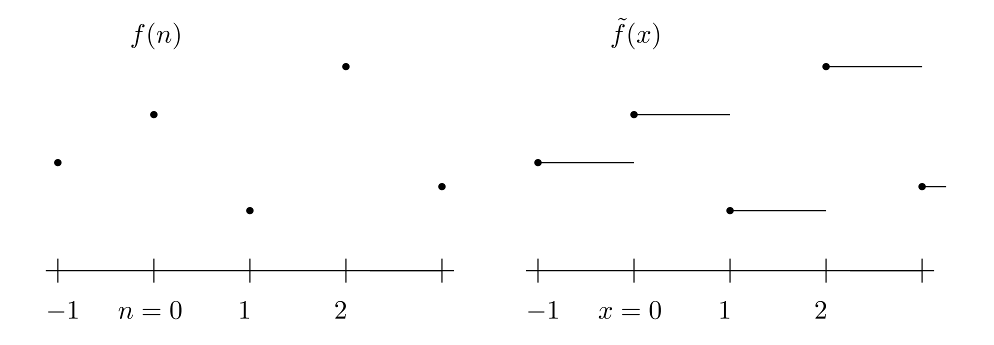


Figure 2. Extension of f to $\mathbb{R}$

Similarly, if $E \subset \mathbb{Z}$, denote by $\tilde{E}$ the set in $\mathbb{R}$ given by $\tilde{E} = \bigcup_{n \in E} [n, n + 1)$. Note that as a result of these definitions we have $m(\tilde{E}) = \#(E)$ and $\int_{\mathbb{R}} \tilde{f}(x)\, dx = \sum_{n \in \mathbb{Z}} f(n)$, and thus $\|\tilde{f}\|_{L^1(\mathbb{R})} = \|f\|_{L^1(\mathbb{Z})}$. Here m is the Lebesgue measure on $\mathbb{R}$, and $\#$ is the counting measure on $\mathbb{Z}$. Note also

that

$$\sum_{k=0}^{m-1} f(n+k) = \int_0^m \tilde{f}(n+t)\,dt.$$

However, because $\int_0^m \tilde{f}(n+t)\,dt \leq \int_{-1}^m \tilde{f}(x+t)\,dt$ whenever $x \in [n, n+1)$, we see that

$$\frac{1}{m}\sum_{k=0}^{m-1} f(n+k) \leq \left(\frac{m+1}{m}\right)\frac{1}{m+1}\int_{-1}^m \tilde{f}(x+t)\,dt \qquad \text{if } x \in [n, n+1).$$

Taking the supremum over all $m \geq 1$ in the above and noting that $(m+1)/m \leq 2$, we obtain

$$f^*(n) \leq 2(\tilde{f})^*(x) \qquad \text{whenever } x \in [n, n+1). \tag{25}$$

To be clear about the notation here: $f^*(n)$ denotes the maximal function of f on $\mathbb{Z}$ defined by (23), with $f(\tau^k(n)) = f(n+k)$, while $(\tilde{f})^*$ is the maximal function as defined in Chapter 3, of the extended function $\tilde{f}$ on $\mathbb{R}$.

By (25)

$$\#(\{n : \ f^*(n) > \alpha\}) \leq m(\{x \in \mathbb{R} : \ (\tilde{f})^*(x) > \alpha/2\}),$$

and thus the latter is majorized by $A'/(\alpha/2)\int \tilde{f}(x)\,dx = 2A'/\alpha\|\tilde{f}\|_{L^1(\mathbb{R})}$, according to the maximal theorem for $\mathbb{R}$. The constant A' that occurs in that theorem (there denoted by A) can be taken to be 3. Hence we have

$$\#(\{n : \ f^*(n) > \alpha\}) \leq \frac{6}{\alpha}\,\|f\|_{L^1(\mathbb{Z})}, \tag{26}$$

since $\|\tilde{f}\|_{L^1(\mathbb{R})} = \|f\|_{L^1(\mathbb{Z})}$. This disposes of the special case when $X = \mathbb{Z}$.

Step 2. The general case.

By a sleight-of-hand we shall "transfer" the result for $\mathbb{Z}$ just proved to the general case. We proceed as follows.

For every positive integer N, we consider the truncated maximal function f_N^* defined as

$$f_N^*(x) = \sup_{1\leq m\leq N} \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x)).$$

Since $\{f_N^*(x)\}$ forms an increasing sequence with N, and $\lim_{N\to\infty} f_N^*(x) = f^*(x)$ for every x, it suffices to show that

$$\mu\{x:\ f_N^*(x) > \alpha\} \leq \frac{A}{\alpha}\ \|f\|_{L^1(X,\mu)}, \tag{27}$$

with constant A independent of N. Letting $N \to \infty$ will then give the desired result.

So in place of f^* we estimate f_N^*, and to simplify our notation we write the latter as f^*, dropping the N subscript. Our argument will compare the maximal function f^* with the special case arising for $\mathbb{Z}$. To clarify the formula below we temporarily adopt the expedient of denoting the second maximal function by $\mathcal{M}(f)$. Thus for a positive function f on $\mathbb{Z}$ we set

$$\mathcal{M}(f)(n) = \sup_{1\leq m} \frac{1}{m}\sum_{k=0}^{m-1} f(n+k).$$

Now starting with a function f on X that is integrable, we define the function F on $X \times \mathbb{Z}$ by

$$F(x,n) = \begin{cases} f(\tau^n(x)) & \text{if } n \geq 0,\\ 0 & \text{if } n < 0.\end{cases}$$

Then

$$A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x)) = \frac{1}{m}\sum_{k=0}^{m-1} F(x,k).$$

In the above we replace x by $\tau^n(x)$; then since $\tau^k(\tau^n(x)) = \tau^{n+k}(x)$, we have

$$A_m(f)(\tau^n(x)) = \frac{1}{m}\sum_{k=0}^{m-1} F(x,n+k).$$

Now we fix a large positive a and set $b = a + N$. We also write F_b for the truncated function on $X \times \mathbb{Z}$ defined by $F_b(x,n) = F(x,n)$ if $n < b$, $F_b(x,n) = 0$ otherwise. We then have

$$A_m(f)(\tau^n(x)) = \frac{1}{m}\sum_{k=0}^{m-1} F_b(x,n+k) \quad \text{if } m \leq N \text{ and } n < a.$$

Thus

$$f^*(\tau^n(x)) \leq \mathcal{M}(F_b)(x,n) \quad \text{if } n < a. \tag{28}$$

(Recall that f^* is actually f_N^*!) This is the comparison of the two maximal functions we wished to obtain. Now set $E_\alpha = \{x : \; f^*(x) > \alpha\}$. Then by the measure-preserving character of τ, $\mu(\{x : \; f^*(\tau^n(x)) > \alpha\}) = \mu(E_\alpha)$. Hence on the product space $X \times \mathbb{Z}$ the product measure $\mu \times \#$ of the set $\{(x, n) \in X \times \mathbb{Z} : \; f^*(\tau^n(x)) > \alpha, \;\; 0 \leq n < a\}$ equals $a\mu(E_\alpha)$. However, because of (28) the $\mu \times \#$ measure of this set is no more than

$$\int_X \#(\{n \in \mathbb{Z} : \; \mathcal{M}(F_b)(x, n) > \alpha\}) \, d\mu.$$

Because of the maximal estimate (26) for $\mathbb{Z}$, we see that the integrand above is no more than

$$\frac{A}{\alpha} \, \|F_b(x, n)\|_{L^1(\mathbb{Z})} = \frac{A}{\alpha} \sum_{n=0}^{b-1} f(\tau^n(x)),$$

with of course $A = 6$.

Hence, integrating this over X and recalling that $\int_X f(\tau^n(x)) \, d\mu = \int_X f(x) \, d\mu$ gives us

$$a\mu(E_\alpha) \leq \frac{A}{\alpha} \, b \, \|f\|_{L^1(X)} = \frac{A}{\alpha} \, (a + N) \, \|f\|_{L^1(X)}.$$

Thus $\mu(E_\alpha) \leq \frac{A}{\alpha} \left(1 + \frac{N}{a}\right) \|f\|_{L^1(X)}$, and letting $a \to \infty$ yields estimate (27). As we have seen, a final limit as $N \to \infty$ then completes the proof.

5.3 Pointwise ergodic theorem

The last of the series of limit theorems we will study is the **pointwise** (or **individual**) **ergodic** theorem, which combines ideas of the first two theorems. At this stage it will be convenient to assume that the measure space (X, μ) is finite; we can then normalize the measure and suppose $\mu(X) = 1$.

Theorem 5.4 *Suppose f is integrable over X. Then for almost every $x \in X$ the averages $A_m(f) = \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$.*

Corollary 5.5 *If we denote this limit by $P'(f)$, we have that*

$$\int_X |P'(f)(x)| \, d\mu(x) \leq \int_X |f(x)| \, d\mu(x).$$

Moreover $P'(f) = P(f)$ whenever $f \in L^2(X, \mu)$.

The idea of the proof is as follows. We first show that $A_m(f)$ converges to a limit almost everywhere for a set of functions f that is dense in $L^1(X,\mu)$. We then use the maximal theorem to show that this implies the conclusion for all integrable functions.

We remark to begin with that because the total measure of X is 1, we have $L^2(X,\mu) \subset L^1(X,\mu)$ and $\|f\|_{L^1} \leq \|f\|_{L^2}$, and moreover $L^2(X,\mu)$ is dense in $L^1(X,\mu)$. In fact, if f belongs to L^1, consider the sequence $\{f_n\}$ defined by $f_n(x) = f(x)$ if $|f(x)| \leq n$, $f_n(x) = 0$ otherwise. Then each f_n is clearly in L^2, while by the dominated convergence theorem $\|f - f_n\|_{L^1} \to 0$.

Now starting with an integrable f and any $\epsilon > 0$ we shall see that we can write $f = F + H$, where $\|H\|_{L^1} < \epsilon$, and $F = F_0 + (1-T)G$, where both F_0 and G belong to L^2, and $T(F_0) = F_0$, with $T(F_0) = F_0(\tau(x))$. To obtain this decomposition of f, we first write $f = f' + h'$, where $f' \in L^2$ and $\|h'\|_{L^1} < \epsilon/2$, which we can do in view of the density of L^2 in L^1 as seen above. Next, since the subspaces S and $\overline{S_1}$ of Lemma 5.2 are orthogonal complements in L^2, we can find $F_0 \in S$, $F_1 \in S_1$, such that $f' = F_0 + F_1 + h$ with $\|h\|_{L^2} < \epsilon/2$. Because $F_1 \in S_1$ is automatically of the form $F_1 = (1-T)G$, we obtain $f = F + H$, with $F = F_0 + (1 - T)G$ and $H = h + h'$. Thus $\|H\|_{L^1} \leq \|h\|_{L^1} + \|h'\|_{L^1}$ and since $\|h\|_{L^1} \leq \|h\|_{L^2} < \epsilon/2$ we have achieved our desired decomposition of f.

Now $A_m(F) = A_m(F_0) + A_m((1-T)G) = F_0 + \frac{1}{m}(1 - T^m(G))$, as we have already seen in the proof of Theorem 5.1. Note that $\frac{1}{m}T^m(G) = \frac{1}{m}G(\tau^m(x))$ converges to zero as $m \to \infty$ for almost every $x \in X$. Indeed, the series $\sum_{m=1}^{\infty} \frac{1}{m^2}(G(\tau^m(x)))^2$ converges almost everywhere by the monotone convergence theorem, since its integral over X is

$$\sum_{m=1}^{\infty} \frac{1}{m^2}\|T^m G\|_{L^2}^2 = \|G\|_{L^2}^2 \sum_{m=1}^{\infty} \frac{1}{m^2},$$

which is finite.

As a result, $A_m(F)(x)$ converges for almost every $x \in X$. Finally, to prove the corresponding convergence for $A_m(f)(x)$, we argue as in Theorem 1.3 in Chapter 3 and set

$$E_\alpha = \{x : \lim_{N\to\infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Then it suffices to see that $\mu(E_\alpha) = 0$ for all $\alpha > 0$. However, since $A_n(f) - A_m(f) = A_n(F) - A_m(F) + A_n(H) - A_m(H)$, and $A_m(F)(x)$ converges almost everywhere as $m \to \infty$, it follows that almost every point

in the set E_α is contained in E'_α, where

$$E'_\alpha = \{x : \sup_{n,m\geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\},$$

and thus $\mu(E_\alpha) \leq \mu(E'_\alpha) \leq \mu(\{x : 2\sup_m |A_m(H)(x)| > \alpha\})$. The last quantity is majorized by $A/(\alpha/2)\|H\|_{L^1} \leq 2\epsilon A/\alpha$ by Theorem 5.3. Since ϵ was arbitrary we see that $\mu(E_\alpha) = 0$, and hence $A_m(f)(x)$ is a Cauchy sequence for almost every x, and the theorem is proved.

To establish the corollary, observe that if $f \in L^2(X)$, we know by Theorem 5.1 that $A_m(f)$ converges to $P(f)$ in the L^2-norm, and hence a subsequence converges almost everywhere to that limit, showing that $P(f) = P'(f)$ in that case. Next, for any f that is merely integrable, we have

$$\int_X |A_m(f)|\,dx \leq \frac{1}{m}\sum_{k=0}^{m-1}\int_X |f(\tau^k(x))|\,d\mu(x) = \int_X |f(x)|\,d\mu(x),$$

and thus since $A_m(f) \to P'(f)$ almost everywhere, we get by Fatou's lemma that $\int_X |P'(f)(x)|\,d\mu(x) \leq \int_X |f(x)|\,d\mu(x)$. With this the corollary is also proved.

It can be shown that the conclusions of the theorem and corollary are still valid if we drop the assumption that the space X has finite measure. The modifications of the argument needed to obtain this more general conclusion are outlined in Exercise 26.

5.4 Ergodic measure-preserving transformations

The adjective "ergodic" is commonly applied to the three limit theorems proved above. It also has a related but separate usage describing an important class of transformations of the space X.

We say that a measure-preserving transformation τ of X is **ergodic** if whenever E is a measurable set that is "invariant," that is, E and $\tau^{-1}(E)$ differ by sets of measure zero, then either E or E^c has measure zero.

There is a useful rephrasing of this condition of ergodicity. Expanding the definition used in Section 5.1 we say that a measurable function f is **invariant** if $f(x) = f(\tau(x))$ for a.e. $x \in X$. Then τ is ergodic exactly when the only invariant functions are equivalent to constants. In fact, let τ be an ergodic transformation, and assume that f is a real-valued invariant function. Then each of the sets $E_a = \{x : f(x) > a\}$ is invariant, hence $\mu(E_a) = 0$ or $\mu(E_a^c) = 0$ for each a. However, if f is not equivalent

to a constant, then both $\mu(E_a)$ and $\mu(E_a^c)$ must have strictly positive measure for some a. In the converse direction we merely need to note that if all characteristic functions of measurable sets that are invariant must be constants, then τ is ergodic.

The following result subsumes the conclusion of Theorem 5.4 for ergodic transformations. We keep to the assumption of that theorem that the underlying space X has measure equal to 1.

Corollary 5.6 *Suppose τ is an ergodic measure-preserving transformation. For any integrable function f we have*

$$\frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x)) \quad \textit{converges to} \quad \int_X f\, d\mu \quad \textit{for a.e. } x \in X \textit{ as } m \to \infty.$$

The result has the interpretation that the "time average" of f equals its "space average."

Proof. By Theorem 5.1 we know that the averages $A_m(f)$ converge to $P(f)$, whenever $f \in L^2$, where P is the orthogonal projection on the subspace of invariant vectors. Since in this case the invariant vectors form a one-dimensional space spanned by the constant functions, we observe that $P(f) = 1(f,1) = \int_X f\, d\mu$, where 1 designates the function identically equal to 1 on X. To verify this, note that P is the identity on constants and annihilates all functions orthogonal to constants. Next we write any $f \in L^1$ as $g + h$, where $g \in L^2$ and $\|h\|_{L^1} < \epsilon$. Then $P'(f) = P'(g) + P'(h)$. However, we also know that $P'(g) = P(g)$, and $\|P'(h)\| \leq \|h\|_{L^1} < \epsilon$ by the corollary to Theorem 5.4. Thus

$$P'(f) - \int_X f\, d\mu = \int_X (g - f)\, d\mu + P'(h)$$

yields that $\|P'(f) - \int_X f\, d\mu\|_{L^1} \leq \|g - f\|_{L^1} + \epsilon < 2\epsilon$. This shows that $P'(f)$ is the constant $\int_X f\, d\mu$ and the assertion is proved.

We shall now elaborate on the nature of ergodicity and illustrate its thrust in terms of several examples.

a) Rotations of the circle

Here we take up the example described in (iii) at the beginning of Section 5*. On the unit circle $\mathbb{R}/\mathbb{Z}$ with the induced Lebesgue measure, we consider the action τ given by $x \mapsto x + \alpha \mod 1$. The result is

- *The mapping τ is ergodic if and only if α is irrational.*

To begin with, if α is irrational we know by the equidistribution theorem that

$$\frac{1}{n}\sum_{k=0}^{n-1} f(x+k\alpha) \to \int_0^1 f(x)\,dx \qquad \text{as } n\to\infty \tag{29}$$

for every x if f is continuous on $[0,1]$ and periodic ($f(0)=f(1)$). The argument used to prove this goes as follows.[2] First we verify that (29) holds whenever $f(x)=e^{2\pi inx}$, $n\in\mathbb{Z}$, by considering the cases $n=0$ and $n\neq 0$ separately. It then follows that (29) is valid for any trigonometric polynomial (a finite linear combination of these exponentials). Finally, any continuous and periodic function can be uniformly approximated by trigonometric polynomials, so (29) goes over to the general case.

Now if P is the projection on invariant L^2-functions, then Theorem 5.1 and (29) show that P projects onto the constants, when restricted to the continuous periodic functions. Since this subspace is dense in L^2, we see that P still projects all of L^2 on constants; hence the invariant L^2-functions are constants and thus τ is ergodic.

On the other hand, suppose $\alpha = p/q$. Choose any set $E_0\subset(0,1/q)$, so that $0<m(E_0)<1/q$, and let E denote the disjoint union $\bigcup_{r=0}^{q-1}(E_0 + r/q)$. Then clearly E is invariant under $\tau: x\mapsto x+p/q$, and $0<m(E)=qm(E_0)<1$; thus τ is not ergodic.

The property (29) we used, which involves the existence of the limit at all points, is actually stronger than ergodicity: it implies that the measure $d\mu = dx$ is **uniquely ergodic** for this mapping τ. That means that if ν is any measure on the Borel sets of X preserved by τ and $\nu(X)=1$, then ν must equal μ.

To see that this so in the present case, let P_ν be the orthogonal projection guaranteed by Theorem 5.1, on the space $L^2(X,\nu)$. Then (29) shows again that the range of P_ν on the continuous functions, and then on all of $L^2(X,\nu)$, is the subspace of constants, and thus $P_\nu(f)=\int_0^1 f\,d\nu$.

This means also that $\int_0^1 f(x)\,dx = \int_0^1 f\,d\nu$ whenever f is continuous and periodic. By a simple limiting argument we then get that the measure $dx = d\mu$ and ν agree on all open intervals, and thus on all open sets. As we have seen, this then proves that the two measures are then identical.

In general, uniquely ergodic measure-preserving transformations are ergodic, but the converse need not be true, as we shall see below.

b) The doubling mapping

[2] See also Section 2, Chapter 4 in Book I.

We now consider the mapping $x \mapsto 2x \mod 1$ for $x \in (0,1]$, with μ Lebesgue measure, that arose in example (iv) at the beginning of Section 5*. We shall prove that τ is ergodic and in fact satisfies a different and stronger property called mixing.[3] It is defined as follows.

If τ is a measure-preserving transformation on the space (X, μ), it is said to be **mixing** if whenever E and F are a pair of measurable subsets then

$$\mu(\tau^{-n}(E) \cap F) \to \mu(E)\mu(F) \quad \text{as } n \to \infty. \tag{30}$$

The meaning of (30) can be understood as follows. In probability theory one often encounters a "universe" of possible events to which probabilities are assigned. These events are represented as measurable subsets E of some space (X, μ) with $\mu(X) = 1$. The probability of each event is then $\mu(E)$. Two events E and F are "independent" if the probability that they both occur is the product of their separate probabilities, that is, $\mu(E \cap F) = \mu(E)\mu(F)$. The assertion (30) of mixing is then that in the limit as time n tends to infinity, the sets $\tau^{-n}(E)$ and F are asymptotically independent, whatever the choices of E and F.

We shall next observe that the mixing condition is implied by the seemingly stronger condition

$$(T^n f, g) \to (f, 1)(1, g) \quad \text{as } n \to \infty, \tag{31}$$

where $T^n(f)(x) = f(\tau^n(x))$ whenever f and g belong to $L^2(X, \mu)$. This implication follows immediately upon taking $f = \chi_E$ and $g = \chi_F$. The converse is also true, but we leave its proof as an exercise to the reader.

We now remark that the mixing condition implies the ergodicity of τ. Indeed, by (31)

$$(A_n(f), g) = \frac{1}{n}\sum_{k=0}^{n-1}(T^k f, g) \quad \text{converges to } (f, 1)(1, g).$$

This means $(P(f), g) = (f, 1)(1, g)$, and hence $P(f)$ is orthogonal to all g that are orthogonal to constants. This of course means that P is the orthogonal projection on constants, and hence τ is ergodic.

We next observe that the doubling map is mixing. Indeed, if $f(x) = e^{2\pi imx}$, $g(x) = e^{2\pi ikx}$, then $(f, 1)(1, g) = 0$, unless both m and k are 0, in which case this product equals 1. However, in this case $(T^n f, g) = \int_0^1 e^{2\pi im2^n x} e^{-2\pi ikx}\, dx$, and this vanishes for sufficiently large n, unless

[3] This property is often referred to as a "strongly mixing" to distinguish it from still another kind of ergodicity called "weakly mixing."

both m and k are 0, in which case the integral equals 1. Thus (31) holds for all exponentials $f(x) = e^{2\pi imx}$, $g(x) = e^{2\pi ikx}$, and therefore by linearity for all trigonometric polynomials f and g. It is from there an easy step to use the completeness in Chapter 4 to pass to all f and g in $L^2((0,1])$ by approximating these functions in the L^2-norm by trigonometric polynomials.

Let us observe that the action of rotations $\tau : x \mapsto x + \alpha$ of the unit circle for irrational α, although ergodic, is not mixing. Indeed, if we take $f(x) = g(x) = e^{2\pi imx}$, $m \neq 0$, then $(T^n f, g) = e^{2\pi inm\alpha}(f, g) = e^{2\pi inm\alpha}$, while $(f, 1) = (1, g) = 0$; thus $(T^n f, g)$ does not converge to $(f, 1)(1, g)$ as $n \to \infty$.

Finally, we note that the doubling map $\tau : x \mapsto 2x \mod 1$ on $(0, 1]$ is not uniquely ergodic. Besides the Lesbesgue measure, the measure ν with $\nu\{1\} = 1$ but $\nu(E) = 0$ if $1 \notin E$ is also preserved by τ.

Further examples of ergodic transformations are given below.

6* Appendix: the spectral theorem

The purpose of this appendix is to present an outline of the proof of the spectral theorem for bounded symmetric operators on a Hilbert space. Details that are not central to the proof of the theorem will be left to the interested reader to fill in. The theorem provides an interesting application of the ideas related to the Lebesgue-Stieltjes integrals that are treated in this chapter.

6.1 Statement of the theorem

A basic notion is that of a **spectral resolution** (or **spectral family**) on a Hilbert space $\mathcal{H}$. This is a function $\lambda \mapsto E(\lambda)$ from $\mathbb{R}$ to orthogonal projections on $\mathcal{H}$ that satisfies the following:

(i) $E(\lambda)$ is increasing in the sense that $\|E(\lambda)f\|$ is an increasing function of λ for every $f \in \mathcal{H}$.

(ii) There is an interval $[a, b]$ such that $E(\lambda) = 0$ if $\lambda < a$, and $E(\lambda) = I$ if $\lambda \geq b$. Here I denotes the identity operator on $\mathcal{H}$.

(iii) $E(\lambda)$ is right-continuous, that is, for every λ one has

$$\lim_{\substack{\mu \to \lambda \\ \mu > \lambda}} E(\mu)f = E(\lambda)f \qquad \text{for every } f \in \mathcal{H}.$$

Observe that property (i) is equivalent with each of the following three assertions (holding for all pairs λ, μ with $\mu > \lambda$): (a) the range of $E(\mu)$ contains the range of $E(\lambda)$; (b) $E(\mu)E(\lambda) = E(\lambda)$; (c) $E(\mu) - E(\lambda)$ is an orthogonal projection.

Now given a spectral resolution $\{E(\lambda)\}$ and an element $f \in \mathcal{H}$, note that the function $\lambda \mapsto (E(\lambda)f, f) = \|E(\lambda)f\|^2$ is also increasing. As a result, the polarization identity (see Section 5 in Chapter 4) shows that for every pair $f, g \in \mathcal{H}$,

the function $F(\lambda) = (E(\lambda)f, g)$ is of bounded variation, and is moreover right-continuous. With these two observations we can now state the main result.

Theorem 6.1 *Suppose T is a bounded symmetric operator on a Hilbert space $\mathcal{H}$. Then there exists a spectral resolution $\{E(\lambda)\}$ such that*

$$T = \int_{a^-}^{b} \lambda \, dE(\lambda)$$

in the sense that for every $f, g \in \mathcal{H}$

$$(Tf, g) = \int_{a^-}^{b} \lambda \, d(E(\lambda)f, g) = \int_{a^-}^{b} \lambda \, dF(\lambda). \tag{32}$$

The integral on the right-hand side is taken in the Lebesgue-Stieltjes sense, as in (iii) and (iv) of Section 3.3.

The result encompasses the spectral theorem for compact symmetric operators T in the following sense. Let $\{\varphi_k\}$ be an orthonormal basis of eigenvectors of T with corresponding eigenvalues λ_k, as guaranteed by Theorem 6.2 in Chapter 4. In this case, we take the spectral resolution to be defined via this orthogonal expansion by

$$E(\lambda)f \sim \sum_{\lambda_k \leq \lambda} (f, \varphi_k)\varphi_k,$$

and one easily verifies that it satisfies conditions (i), (ii) and (iii) above. We also note that $\|E(\lambda)f\|^2 = \sum_{\lambda_k \leq \lambda} |(f, \varphi_k)|^2$, and thus $F(\lambda) = (E(\lambda)f, g)$ is a pure jump function as in Section 3.3 in Chapter 3.

6.2 Positive operators

The proof of the theorem depends on the concept of positivity of operators. We say that T is **positive**, written as $T \geq 0$, if T is symmetric and $(Tf, f) \geq 0$ for all $f \in \mathcal{H}$. (Note that (Tf, f) is automatically real if T is symmetric.) One then writes $T_1 \geq T_2$ to mean that $T_1 - T_2 \geq 0$. Note that for two orthogonal projections we have $E_2 \geq E_1$ if and only if $\|E_2 f\| \geq \|E_1 f\|$ for all $f \in \mathcal{H}$, and that is then equivalent with the corresponding properties (a)−(c) described above. Notice also that if S is symmetric, then $S^2 = T$ is positive. Now for T symmetric, let us write

$$a = \min(Tf, f) \quad \text{and} \quad b = \max(Tf, f) \quad \text{for } \|f\| \leq 1. \tag{33}$$

Proposition 6.2 *Suppose T is symmetric. Then $\|T\| \leq M$ if and only if $-MI \leq T \leq MI$. As a result, $\|T\| = \max(|a|, |b|)$.*

This is a consequence of (7) in Chapter 4.

Proposition 6.3 *Suppose T is positive. Then there exists a symmetric operator S (which can be written as $T^{1/2}$) such that $S^2 = T$ and S commutes with every operator that commutes with T.*

The last assertion means that if for some operator A we have $AT = TA$, then $AS = SA$.

The existence of S is seen as follows. After multiplying by a suitable positive scalar, we may assume that $\|T\| \leq 1$. Consider the binomial expansion of $(1-t)^{1/2}$, given by $(1-t)^{1/2} = \sum_{k=0}^{\infty} b_k t^k$, for $|t| < 1$. The relevant fact that is needed here is that the b_k are real and $\sum_{k=0}^{\infty} |b_k| < \infty$. Indeed, by direct calculation of the power series expansion of $(1-t)^{1/2}$ we find that $b_0 = 1$, $b_1 = -1/2$, $b_2 = -1/8$, and more generally, $b_k = -1/2 \cdot 1/2 \cdots (k-3/2)/k!$, if $k \geq 2$, from which it follows that $b_k = O(k^{-3/2})$. Or more simply, since $b_k < 0$ when $k \geq 1$, if we let $t \to 1$ in the definition, we see that $-\sum_{k=1}^{\infty} b_k = 1$ and so $\sum_{k=0}^{\infty} |b_k| = 2$.

Now let $s_n(t)$ denote the polynomial $\sum_{k=0}^{n} b_k t^k$. Then the polynomial

$$s_n^2(t) - (1-t) = \sum_{k=0}^{2n} c_k^n t^k \tag{34}$$

has the property that $\sum_{k=0}^{2n} |c_k^n| \to 0$ as $n \to \infty$. In fact, $s_n(t) = (1-t)^{1/2} - r_n(t)$, with $r_n(t) = \sum_{k=n+1}^{\infty} b_k t^k$, so $s_n^2(t) - (1-t) = -r_n^2(t) - 2s_n(t)r_n(t)$. Now the left-hand side is clearly a polynomial of degree $\leq 2n$, and so comparing coefficients with those on the right-hand side shows that the c_k^n are majorized by $3\sum_{j>n} |b_j|\,|b_{k-j}|$. From this it is immediate that $\sum_k |c_k^n| = O(\sum_{j>n} |b_j|) \to 0$ as $n \to \infty$, as asserted.

To apply this, set $T_1 = I - T$; then $0 \leq T_1 \leq I$, and thus $\|T_1\| \leq 1$, by Proposition 6.2. Let $S_n = s_n(T_1) = \sum_{k=0}^{n} b_k T_1^k$, with $T_1^0 = I$. Then in terms of operator norms, $\|S_n - S_m\| \leq \sum_{k \geq \min(n,m)} |b_k| \to 0$ as $n, m \to \infty$, because $\|T_1^k\| \leq \|T_1\|^k \leq 1$. Hence S_n converges to some operator S. Clearly S_n is symmetric for each n, and thus S is also symmetric. Moreover, by (34), $S_n^2 - T = \sum_{k=0}^{2n} c_k^n T_1^k$, therefore $\|S_n^2 - T\| \leq \sum |c_k^n| \to 0$ as $n \to \infty$, which implies that $S^2 = T$. Finally, if A commutes with T it clearly commutes with every polynomial in T, hence with S_n, and thus with S. The proof of the proposition is therefore complete.

Proposition 6.4 *If T_1 and T_2 are positive operators that commute, then T_1T_2 is also positive.*

Indeed, if S is a square root of T_1 given in the previous proposition, then $T_1T_2 = SST_2 = ST_2S$, and hence $(T_1T_2f, f) = (ST_2Sf, f) = (T_2Sf, Sf)$, since S is symmetric, and thus the last term is positive.

Proposition 6.5 *Suppose T is symmetric and a and b are given by* (33). *If $p(t) = \sum_{k=0}^{n} c_k t^k$ is a real polynomial which is positive for $t \in [a,b]$, then the operator $p(T) = \sum_{k=0}^{n} c_k T^k$ is positive.*

To see this, write $p(t) = c\prod_j (t-\rho_j)\prod_k(\rho_k' - t)\prod_\ell((t-\mu_\ell)^2 + \nu_\ell)$, where c is positive and the third factor corresponds to the non-real roots of $p(t)$ (arising in conjugate pairs), and the real roots of $p(t)$ lying in (a,b) which are necessarily of even order. The first factor contains the real roots ρ_j with $\rho_j \leq a$, and the second factor the real roots ρ_k' with $\rho_k' \geq b$. Since each of the factors $T - \rho_j I$, $\rho_j' I - T$ and $(T - \mu_\ell I)^2 + \nu_\ell^2 I$ is positive and these commute, the desired conclusion follows from the previous proposition.

Corollary 6.6 *If $p(t)$ is a real polynomial, then*

$$\|p(T)\| \leq \sup_{t\in[a,b]} |p(t)|.$$

This is an immediate consequence using Proposition 6.2, since $-M \leq p(t) \leq M$, where $M = \sup_{t\in[a,b]} |p(t)|$, and thus $-MI \leq p(T) \leq MI$.

Proposition 6.7 *Suppose $\{T_n\}$ is a sequence of positive operators that satisfy $T_n \geq T_{n+1}$ for all n. Then there is a positive operator T, such that $T_n f \to Tf$ as $n \to \infty$ for every $f \in \mathcal{H}$.*

Proof. We note that for each fixed $f \in \mathcal{H}$ the sequence of positive numbers $(T_n f, f)$ is decreasing and hence convergent. Now observe that for any positive operator S with $\|S\| \leq M$ we have

$$\|S(f)\|^2 \leq (Sf,f)^{1/2} M^{3/2} \|f\|. \tag{35}$$

In fact, the quadratic function $(S(tI+S)f, (tI+S)f) = t^2(Sf,f) + 2t(Sf,Sf) + (S^2f, Sf)$ is positive for all real t. Hence its discriminant is negative, that is, $\|S(f)\|^4 \leq (Sf,f)(S^2f,Sf)$, and (35) follows. We apply this to $S = T_n - T_m$ with $n \leq m$; then $\|T_n - T_m\| \leq \|T_n\| \leq \|T_1\| = M$, and since $((T_n - T_m)f, f) \to 0$ as $n, m \to \infty$ we see that $\|T_n f - T_m f\| \to 0$ as $n, m \to \infty$. Thus $\lim_{n\to\infty} T_n(f) = T(f)$ exists, and T is also clearly positive.

6.3 Proof of the theorem

Starting with a given symmetric operator T, and with a, b given by (33), we shall now exploit further the idea of associating to each suitable function Φ on $[a, b]$ a symmetric operator $\Phi(T)$. We do this in increasing order of generality. First, if Φ is a real polynomial $\sum_{k=0}^n c_k t^k$, then, as before, $\Phi(T)$ is defined as $\sum_{k=0}^n c_k T^k$. Notice that this association is a homomorphism: if $\Phi = \Phi_1 + \Phi_2$, then $\Phi(T) = \Phi_1(T) + \Phi_2(T)$; also if $\Phi = \Phi_1 \cdot \Phi_2$, then $\Phi(T) = \Phi_1(T) \cdot \Phi_2(T)$. Moreover, since Φ is real (and the c_k are real), $\Phi(T)$ is symmetric.

Next, because every real-valued continuous function Φ on $[a, b]$ can be approximated uniformly by polynomials p_n (see, for instance, Section 1.8, Chapter 5 of Book I), we see by Corollary 6.6 that the sequence $p_n(T)$ converges, in the norm of operators, to a limit which we call $\Phi(T)$, and moreover this limit does not depend on the sequence of polynomials approximating Φ. Also, $\Phi(T)$ is automatically a symmetric operator. If $\Phi(t) \geq 0$ on $[a, b]$ we can always take the approximating sequence to be positive on $[a, b]$, and as a result $\Phi(T) \geq 0$.

Finally, we define $\Phi(T)$ whenever Φ arises as a limit, $\Phi(t) = \lim_{n\to\infty} \Phi_n(t)$, where $\{\Phi_n(t)\}$ is a decreasing sequence of positive continuous functions on $[a, b]$. In fact, by Proposition 6.7 the limit $\lim_{n\to\infty} \Phi_n(T)$ exists by what we have established above for Φ_n. To show that this limit is independent of the sequence $\{\Phi_n\}$ and thus that $\Phi(t)$ is well-defined as the limit above, let $\{\Phi_n'\}$ be another sequence of decreasing continuous functions converging to Φ. Then whenever $\epsilon > 0$ is given and k is fixed, $\Phi_n'(t) \leq \Phi_k(t) + \epsilon$ for all n sufficiently large. Thus $\Phi_n'(T) \leq \Phi_k(T) + \epsilon I$ for these n, and passing to the limit first in n, then in k, and then with $\epsilon \to 0$, we get

$\lim_{n\to\infty} \Phi'_n(T) \leq \lim_{k\to\infty} \Phi_k(T)$. By symmetry, the reverse inequality holds, and the two limits are the same. Note also that for a pair of these limiting functions, if $\Phi_1(t) \leq \Phi_2(t)$ for $t \in [a, b]$, then $\Phi_1(T) \leq \Phi_2(T)$.

The basic functions Φ, $\Phi = \varphi^\lambda$, that give us the spectral resolution are defined for each real λ by

$$\varphi^\lambda(t) = 1 \quad \text{if } t \leq \lambda \quad \text{and} \quad \varphi^\lambda(t) = 0 \quad \text{if } \lambda < t.$$

We note that $\varphi^\lambda(t) = \lim \varphi_n^\lambda(t)$, where $\varphi_n^\lambda(t) = 1$ if $t \leq \lambda$, $\varphi_n^\lambda(t) = 0$ if $t \geq \lambda + 1/n$, and $\varphi_n^\lambda(t)$ is linear for $t \in [\lambda, \lambda + 1/n]$. Thus each $\varphi^\lambda(t)$ is a limit of a decreasing sequence of continuous functions. In accordance with the above we set

$$E(\lambda) = \varphi^\lambda(T).$$

Since $\lim_{n\to\infty} \varphi_n^{\lambda_1}(t)\varphi_n^{\lambda_2}(t) = \varphi_n^{\lambda_1}(t)$ whenever $\lambda_1 \leq \lambda_2$, we see that $E(\lambda_1)E(\lambda_2) = E(\lambda_1)$. Thus $E(\lambda)^2 = E(\lambda)$ for every λ, and because $E(\lambda)$ is symmetric it is therefore an orthogonal projection. Moreover, for every $f \in \mathcal{H}$

$$\|E(\lambda_1)f\| = \|E(\lambda_1)E(\lambda_2)f\| \leq \|E(\lambda_2)f\|,$$

thus $E(\lambda)$ is increasing. Clearly $E(\lambda) = 0$ if $\lambda < a$, since for those λ, $\varphi^\lambda(t) = 0$ on $[a, b]$. Similarly, $E(\lambda) = I$ for $\lambda \geq b$.

Next we note that $E(\lambda)$ is right-continuous. In fact, fix $f \in \mathcal{H}$ and $\epsilon > 0$. Then for some n, which we now keep fixed, $\|E(\lambda)f - \varphi_n^\lambda(T)f\| < \epsilon$. However, $\varphi_n^\mu(t)$ converges to $\varphi_n^\lambda(t)$ uniformly in t as $\mu \to \lambda$. Hence $\sup_t |\varphi_n^\mu(t) - \varphi_n^\lambda(t)| < \epsilon$, if $|\mu - \lambda| < \delta$, for an appropriate δ. Thus by the corollary $\|\varphi_n^\mu(T) - \varphi_n^\lambda(T)\| < \epsilon$ and therefore $\|E(\lambda)f - \varphi_n^\mu(T)\| < 2\epsilon$. Now with $\mu \geq \lambda$ we have that $E(\mu)E(\lambda) = E(\lambda)$ and $E(\mu)\varphi_n^\mu(T) = E(\mu)$. As a result $\|E(\lambda)f - E(\mu)f\| < 2\epsilon$, if $\lambda \leq \mu \leq \lambda + \delta$. Since ϵ was arbitrary, the right continuity is established.

Finally we verify the spectral representation (32). Let $a = \lambda_0 < \lambda_1 < \cdots < \lambda_k = b$ be any partition of $[a, b]$ for which $\sup_j(\lambda_j - \lambda_{j-1}) < \delta$. Then since

$$t = \sum_{j=1}^{k} t(\varphi^{\lambda_j}(t) - \varphi^{\lambda_{j-1}}(t)) + t\varphi^{\lambda_0}(t)$$

we note that

$$t \leq \sum_{j=1}^{k} \lambda_j(\varphi^{\lambda_j}(t) - \varphi^{\lambda_{j-1}}(t)) + \lambda_0\varphi^{\lambda_0}(t) \leq t + \delta.$$

Applying these functions to the operator T we obtain

$$T \leq \sum_{j=1}^{k} \lambda_j(E(\lambda_j) - E(\lambda_{j-1})) + \lambda_0 E(\lambda_0) \leq T + \delta I,$$

and thus T differs in norm from the sum above by at most δ. As a result

$$\left|(Tf,f) - \sum_{j=1}^{k} \lambda_j \int_{(\lambda_{j-1},\lambda_j]} d(E(\lambda)f,f) - \lambda_0(E(\lambda_0)f,f)\right| \leq \delta\|f\|^2.$$

But as we vary the partitions of $[a,b]$, letting their meshes δ tend to zero, the above sum tends to $\int_{a-}^{b} \lambda\, d(E(\lambda)f,f)$. Therefore $(Tf,f) = \int_{a-}^{b} \lambda\, d(E(\lambda)f,f)$, and the polarization identity gives (32).

A similar argument shows that if Φ is continuous on $[a,b]$, then the operator $\Phi(T)$ has an analogous spectral representation

$$(36) \qquad (\Phi(T)f,g) = \int_{a-}^{b} \Phi(\lambda)\, d(E(\lambda)f,g).$$

This is because $|\Phi(t) - \sum_{j=1}^{k} \Phi(\lambda_j)(\varphi^{\lambda_j}(t) - \varphi^{\lambda_{j-1}}(t)) - \Phi(\lambda_0)\varphi^{\lambda_0}(t)| < \delta'$, where $\delta' = \sup_{|t-t'|\leq\delta} |\Phi(t) - \Phi(t')|$, which tends to zero as $\delta \to 0$.

This representation also extends to continuous Φ that are complex-valued (by considering the real and imaginary parts separately) or for Φ that are limits of decreasing pointwise continuous functions.

6.4 Spectrum

We say that a bounded operator S on $\mathcal{H}$ is **invertible** if S is a bijection of $\mathcal{H}$ and its inverse, S^{-1}, is also bounded. Note that S^{-1} satisfies $S^{-1}S = SS^{-1} = I$. The **spectrum** of S, denoted by $\sigma(S)$, is the set of complex numbers z for which $S - zI$ is *not* invertible.

Proposition 6.8 *If T is symmetric, then $\sigma(T)$ is a closed subset of the interval $[a,b]$ given by* (33).

Note that if $z \notin [a,b]$, the function $\Phi(t) = (t-z)^{-1}$ is continuous on $[a,b]$ and $\Phi(T)(T-zI) = (T-zI)\Phi(T) = I$, so $\Phi(T)$ is the inverse of $T - zI$. Now suppose $T_0 = T - \lambda_0 I$ is invertible. Then we claim that $T_0 - \epsilon I$ is invertible for all (complex) ϵ that are sufficiently small. This will prove that the complement of $\sigma(T)$ is open. Indeed, $T_0 - \epsilon I = T_0(I - \epsilon T_0^{-1})$, and we can invert the operator $(I - \epsilon T_0^{-1})$ (formally) by writing its inverse as a sum

$$\sum_{n=0}^{\infty} \epsilon^n (T_0^{-1})^{n+1}.$$

Since $\sum_{n=0}^{\infty} \|\epsilon^n (T_0^{-1})^{n+1}\| \leq \sum |\epsilon|^n \|T_0^{-1}\|^{n+1}$, the series converges when $|\epsilon| < \|T_0^{-1}\|^{-1}$, and the sum is majorized by

$$(37) \qquad \|T_0^{-1}\| \frac{1}{1 - |\epsilon|\|T_0^{-1}\|}.$$

Thus we can define the operator $(T_0 - \epsilon I)^{-1}$ as $\lim_{N\to\infty} T_0^{-1} \sum_{n=0}^{N} \epsilon^n (T_0^{-1})^{n+1}$, and it gives the desired inverse, as is easily verified.

Our last assertion connects the spectrum $\sigma(T)$ with the spectral resolution $\{E(\lambda)\}$.

Proposition 6.9 *For each $f \in \mathcal{H}$, the Lebesgue-Stieltjes measure corresponding to $F(\lambda) = (E(\lambda)f, f)$ is supported on $\sigma(T)$.*

To put it another way, $F(\lambda)$ is constant on each open interval of the complement of $\sigma(T)$.

To prove this, let J be one of the open intervals in the complement of $\sigma(T)$, $x_0 \in J$, and J_0 the sub-interval centered at x_0 of length 2ϵ, with $\epsilon < \|(T - x_0 I)^{-1}\|$. First note that if z has non-vanishing imaginary part then $(T - zI)^{-1}$ is given by $\Phi_z(T)$, with $\Phi_z(t) = (t - z)^{-1}$. Hence $(T - zI)^{-1}(T - \overline{z}I)^{-1}$ is given by $\Psi_z(T)$, with $\Psi_z(t) = 1/|t - z|^2$. Therefore by the estimate given in (37) and the representation (36) applied to $\Phi = \Psi_z$, we obtain

$$\int \frac{dF(\lambda)}{|\lambda - z|^2} \leq A',$$

as long as z is complex and $|x_0 - z| < \epsilon$. We can therefore obtain the same inequality for x real, $|x_0 - x| < \epsilon$. Now integration in $x \in J_0$ using the fact that $\int_{J_\epsilon} \frac{dx}{|\lambda - x|^2} = \infty$ for every $\lambda \in J_\epsilon$, gives $\int_{J_\epsilon} dF(\lambda) = 0$. Thus $F(\lambda)$ is constant in J_ϵ, but since x_0 was an arbitrary point of J the function $F(\lambda)$ is constant throughout J and the proposition is proved.

7 Exercises

1. Let X be a set and $\mathcal{M}$ a non-empty collection of subsets of X. Prove that if $\mathcal{M}$ is closed under complements and countable unions of disjoint sets, then $\mathcal{M}$ is a σ-algebra.

[Hint: Any countable union of sets can be written as a countable union of disjoint sets.]

2. Let $(X, \mathcal{M}, \mu)$ be a measure space. One can define the **completion** of this space as follows. Let $\overline{\mathcal{M}}$ be the collection of sets of the form $E \cup Z$, where $E \in \mathcal{M}$, and $Z \subset F$ with $F \in \mathcal{M}$ and $\mu(F) = 0$. Also, define $\overline{\mu}(E \cup Z) = \mu(E)$. Then:

(a) $\overline{\mathcal{M}}$ is the smallest σ-algebra containing $\mathcal{M}$ and all subsets of elements of $\mathcal{M}$ of measure zero.

(b) The function $\overline{\mu}$ is a measure on $\overline{\mathcal{M}}$, and this measure is complete.

[Hint: To prove $\overline{\mathcal{M}}$ is a σ-algebra it suffices to see that if $E_1 \subset \overline{\mathcal{M}}$, then $E_1^c \subset \overline{\mathcal{M}}$. Write $E_1 = E \cup Z$ with $Z \subset F$, E and F in $\mathcal{M}$. Then $E_1^c = (E \cup F)^c \cup (F - Z)$.]

3. Consider the exterior Lebesgue measure m_* introduced in Chapter 1. Prove that a set E in $\mathbb{R}^d$ is Carathéodory measurable if and only if E is Lebesgue measurable in the sense of Chapter 1.

[Hint: If E is Lebesgue measurable and A is any set, choose a G_δ set G such that $A \subset G$ and $m_*(A) = m(G)$. Conversely, if E is Carathéodory measurable and $m_*(E) < \infty$, choose a G_δ set G with $E \subset G$ and $m_*(E) = m_*(G)$. Then $G - E$ has exterior measure 0.]

4. Let r be a rotation of $\mathbb{R}^d$. Using the fact that the mapping $x \mapsto r(x)$ preserves Lebesgue measure (see Problem 4 in Chapter 2 and Exercise 26 in Chapter 3), show that it induces a measure-preserving map of the sphere S^{d-1} with its measure $d\sigma$.

A converse is stated in Problem 4.

5. Use the polar coordinate formula to prove the following:

(a) $\int_{\mathbb{R}^d} e^{-\pi|x|^2}\,dx = 1$, when $d = 2$. Deduce from this that the same identity holds for all d.

(b) $\left(\int_0^\infty e^{-\pi r^2} r^{d-1}\,dr\right)\sigma(S^{d-1}) = 1$, and as a result, $\sigma(S^{d-1}) = 2\pi^{d/2}/\Gamma(d/2)$.

(c) If B is the unit ball, $v_d = m(B) = \pi^{d/2}/\Gamma(d/2+1)$, since this quantity equals $\left(\int_0^1 r^{d-1}\,dr\right)\sigma(S^{d-1})$. (See Exercise 14 in Chapter 2.)

6. A version of Green's formula for the unit ball B in $\mathbb{R}^d$ can be stated as follows. Suppose u and v are a pair of functions that are in $C^2(\overline{B})$. Then one has

$$\int_B (v\triangle u - u\triangle v)\,dx = \int_{S^{d-1}} \left(v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n}\right)\,d\sigma.$$

Here S^{d-1} is the unit sphere with $d\sigma$ the measure defined in Section 3.2, and $\partial u/\partial n$, $\partial v/\partial n$ denote the directional derivatives of u and v (respectively) along the inner normals to S^{d-1}.

Show that the above can be derived from Lemma 4.5 of the previous chapter by taking $\eta = \eta_\epsilon^+$ and letting $\epsilon \to 0$.

7. There is an alternate version of the mean-value property given in (21) of Chapter 5. It can be stated as follows. Suppose u is harmonic in Ω, and B is any ball of center x_0 and radius r whose closure is contained in Ω. Then

$$u(x_0) = c\int_{S^{d-1}} u(x_0 + ry)\,d\sigma(y), \qquad \text{with } c^{-1} = \sigma(S^{d-1}).$$

Conversely, a continuous function satisfying this mean-value property is harmonic.

[Hint: This can be proved as a direct consequence of the corresponding result for averages over balls (Theorem 4.27 in Chapter 5), or can be deduced from Exercise 6.]

8. The fact that the Lebesgue measure is uniquely characterized by its translation invariance can be made precise by the following assertion: If μ is a Borel measure on $\mathbb{R}^d$ that is translation-invariant, and is finite on compact sets, then μ is a multiple of Lebesgue measure m. Prove this theorem by proceeding as follows.

(a) Suppose Q_a denotes a translate of the cube $\{x :\ 0 < x_j \le a,\ j = 1, 2, \ldots, d\}$ of side length a. If we let $\mu(Q_1) = c$, then $\mu(Q_{1/n}) = cn^{-d}$ for each integer n.

(b) As a result μ is absolutely continuous with respect to m, and there is a locally integrable function f such that

$$\mu(E) = \int_E f\,dx.$$

(c) By the differentiation theorem (Corollary 1.7 in Chapter 3) it follows that $f(x) = c$ a.e., and hence $\mu = cm$.

[Hint: Q_1 can be written as a disjoint union of n^d translates of $Q_{1/n}$.]

9. Let $C([a,b])$ denote the vector space of continuous functions on the closed and bounded interval $[a,b]$. Suppose we are given a Borel measure μ on this interval, with $\mu([a,b]) < \infty$. Then

$$f \mapsto \ell(f) = \int_a^b f(x)\,d\mu(x)$$

is a linear functional on $C([a,b])$, with ℓ positive in the sense that $\ell(f) \geq 0$ if $f \geq 0$.

Prove that, conversely, for any linear functional ℓ on $C([a,b])$ that is positive in the above sense, there is a unique finite Borel measure μ so that $\ell(f) = \int_a^b f\,d\mu$ for $f \in C([a,b])$.

[Hint: Suppose $a = 0$ and $u \geq 0$. Define $F(u)$ by $F(u) = \lim_{\epsilon \to 0} \ell(f_\epsilon)$, where

$$f_\epsilon(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq u, \\ 0 & \text{for } u + \epsilon \leq x, \end{cases}$$

and f_ϵ is linear between u and $u + \epsilon$. (See Figure 3.) Then F is increasing and right-continuous, and $\ell(f)$ can be written as $\int_a^b f(x)\,dF(x)$ via Theorem 3.5.]

The result also holds if $[a,b]$ is replaced by a closed infinite interval; we then assume that ℓ is defined on the continuous functions of bounded support, and obtain that the resulting μ is finite on all bounded intervals.

A generalization is given in Problem 5.

10. Suppose ν, ν_1, ν_2 are signed measures on $(X, \mathcal{M})$ and μ a (positive) measure on $\mathcal{M}$. Using the symbols $\perp$ and $\ll$ defined in Section 4.2, prove:

(a) If $\nu_1 \perp \mu$ and $\nu_2 \perp \mu$, then $\nu_1 + \nu_2 \perp \mu$.

(b) If $\nu_1 \ll \mu$ and $\nu_2 \ll \mu$, then $\nu_1 + \nu_2 \ll \mu$.

(c) $\nu_1 \perp \nu_2$ implies $|\nu_1| \perp |\nu_2|$.

(d) $\nu \ll |\nu|$.

(e) If $\nu \perp \mu$ and $\nu \ll \mu$, then $\nu = 0$.

11. Suppose that F is an increasing normalized function on $\mathbb{R}$, and let $F = F_A + F_C + F_J$ be the decomposition of F in Exercise 24 in Chapter 3; here F_A is

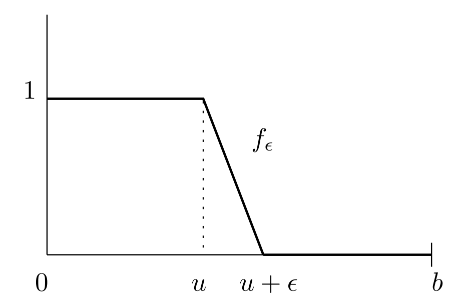


Figure 3. The function f_ϵ in Exercise 9

absolutely continuous, F_C is continuous with $F_C' = 0$ a.e, and F_J is a pure jump function. Let $\mu = \mu_A + \mu_C + \mu_J$ with μ, μ_A, μ_C, and μ_J the Borel measures associated to F, F_A, F_C, and F_J, respectively. Verify that:

(i) μ_A is absolutely continuous with respect to Lebesgue measure and $\mu_A(E) = \int_E F'(x)\,dx$ for every Lebesgue measurable set E.

(ii) As a result, if F is absolutely continuous, then $\int f\,d\mu = \int f\,dF = \int f(x)F'(x)\,dx$ whenever f and fF' are integrable.

(iii) $\mu_C + \mu_J$ and Lebesgue measure are mutually singular.

12. Suppose $\mathbb{R}^d - \{0\}$ is represented as $\mathbb{R}_+ \times S^{d-1}$, with $\mathbb{R}_+ = \{0 < r < \infty\}$. Then every open set in $\mathbb{R}^d - \{0\}$ can be written as a countable union of open rectangles of this product.

[Hint: Consider the countable collection of rectangles of the form

$$\{r_j < r < r_k'\} \times \{\gamma \in S^{d-1} : \ |\gamma - \gamma_\ell| < 1/n\}.$$

Here r_j and r_k' range over all positive rationals, and $\{\gamma_\ell\}$ is a countable dense set of S^{d-1}.]

13. Let m_j be the Lebesgue measure for the space $\mathbb{R}^{d_j}$, $j = 1, 2$. Consider the product $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ ($d = d_1 + d_2$), with m the Lebesgue measure on $\mathbb{R}^d$. Show that m is the completion (in the sense of Exercise 2) of the product measure $m_1 \times m_2$.

14. Suppose $(X_j, \mathcal{M}_j, \mu_j)$, $1 \leq j \leq k$, is a finite collection of measure spaces. Show that parallel with the case $k = 2$ considered in Section 3 one can construct a product measure $\mu_1 \times \mu_2 \times \cdots \times \mu_k$ on $X = X_1 \times X_2 \times \cdots \times X_k$. In fact, for any set $E \subset X$ such that $E = E_1 \times E_2 \times \cdots \times E_k$, with $E_j \subset \mathcal{M}_j$ for all j, define $\mu_0(E) = \prod_{j=1}^k \mu_j(E_j)$. Verify that μ_0 extends to a premeasure on the algebra $\mathcal{A}$ of finite disjoint unions of such sets, and then apply Theorem 1.5.

15. The product theory extends to infinitely many factors, under the requisite assumptions. We consider measure spaces $(X_j, \mathcal{M}_j, \mu_j)$ with $\mu_j(X_j) = 1$ for all but finitely many j. Define a **cylinder set** E as

$$\{x = (x_j),\ x_j \in E_j,\ E_j \in \mathcal{M}_j, \text{but } E_j = X_j \text{ for all but finitely many } j\}.$$

For such a set define $\mu_0(E) = \prod_{j=1}^{\infty} \mu_j(E_j)$. If $\mathcal{A}$ is the algebra generated by the cylinder sets, μ_0 extends to a premeasure on $\mathcal{A}$, and we can apply Theorem 1.5 again.

16. Consider the d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Identify $\mathbb{T}^d$ as $\mathbb{T}^1 \times \cdots \times \mathbb{T}^1$ (d factors) and let μ be the product measure on $\mathbb{T}^d$ given by $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_d$, where μ_j is Lebesgue measure on X_j identified with the circle $\mathbb{T}$. That is, if we represent each point in X_j uniquely as x_j with $0 < x_j \leq 1$, then the measure μ_j is the induced Lebesgue measure on $\mathbb{R}^1$ restricted to $(0, 1]$.

(a) Check that the completion μ is Lebesgue measure induced on the cube $Q = \{x :\ 0 < x_j \leq 1,\ j = 1, \ldots, d\}$.

(b) For each function f on Q let $\tilde{f}$ be its extension to $\mathbb{R}^d$ which is periodic, that is, $\tilde{f}(x + z) = \tilde{f}(x)$ for every $z \in \mathbb{Z}^d$. Then f is measurable on $\mathbb{T}^d$ if and only if $\tilde{f}$ is measurable on $\mathbb{R}^d$, and f is continuous on $\mathbb{T}^d$ if and only if $\tilde{f}$ is continuous on $\mathbb{R}^d$.

(c) Suppose f and g are integrable on $\mathbb{T}^d$. Show that the integral defining $(f * g)(x) = \int_{\mathbb{T}^d} f(x - y)g(y)\, dy$ is finite for a.e. x, that $f * g$ is integrable over $\mathbb{T}^d$, and that $f * g = g * f$.

(d) For any integrable function f on $\mathbb{T}^d$, write

$$f \sim \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n \cdot x}$$

to mean that $a_n = \int_{\mathbb{T}^d} f(x) e^{-2\pi i n \cdot x}\, dx$. Prove that if g is also integrable, and $g \sim \sum_{n \in \mathbb{Z}^d} b_n e^{2\pi i n \cdot x}$, then

$$f * g \sim \sum_{n \in \mathbb{Z}^d} a_n b_n e^{2\pi i n \cdot x}.$$

(e) Verify that $\{e^{2\pi i n \cdot x}\}_{n \in \mathbb{Z}^d}$ is an orthonormal basis for $L^2(\mathbb{T}^d)$. As a result $\|f\|_{L^2(\mathbb{T}^d)} = \sum_{n \in \mathbb{Z}^d} |a_n|^2$.

(f) Let f be any continuous periodic function on $\mathbb{T}^d$. Then f can be uniformly approximated by finite linear combinations of the exponentials $\{e^{2\pi i n \cdot x}\}_{n \in \mathbb{Z}^d}$.

[Hint: For (e), reduce to the case $d = 1$ by Fubini's theorem. To prove (f) let $g(x) = g_\epsilon(x) = \epsilon^{-d}$, if $0 < x_j \leq \epsilon$, $j = 1, \ldots, d$, and $g_\epsilon(x) = 0$ elsewhere in Q. Then $(f * g_\epsilon)(x) \to f(x)$ uniformly as $\epsilon \to 0$. However $(f * g_\epsilon)(x) = \sum a_n b_n e^{2\pi i n x}$ with $b_n = \int_{\mathbb{T}^d} g_\epsilon(x) e^{-2\pi i n \cdot x}\, dx$, and $\sum |a_n b_n| < \infty$.]

17. By reducing to the case $d = 1$, show that each "rotation" $x \mapsto x + \alpha$ of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is measure preserving, for any $\alpha \in \mathbb{R}^d$.

18. Suppose τ is a measure-preserving transformation on a measure space (X, μ) with $\mu(X) = 1$. Recall that a measurable set E is invariant if $\tau^{-1}(E)$ and E differ by a set of measure zero. A sharper notion is to require that $\tau^{-1}(E)$ equal E. Prove that if E is any invariant set, there is a set E' so that $E' = \tau^{-1}(E')$, and E and E' differ by a set of measure zero.

[Hint: Let $E' = \limsup_{n\to\infty}\{\tau^{-n}(E)\} = \bigcap_{n=0}^{\infty}\left(\bigcup_{k\geq n}\tau^{-k}(E)\right)$.]

19. Let τ be a measure-preserving transformation on (X, μ) with $\mu(X) = 1$. Then τ is ergodic if and only if whenever ν is absolutely continuous with respect to μ and ν is invariant (that is, $\nu(\tau^{-1}(E)) = \nu(E)$ for all measurable sets E), then $\nu = c\mu$, with c a constant.

20. Suppose τ is a measure-preserving transformation on (X, μ). If

$$\mu(\tau^{-n}(E) \cap F) \to \mu(E)\mu(F)$$

as $n \to \infty$ for all measurable sets E and F, then $(T^n f, g) \to (f, 1)(1, g)$ whenever $f, g \in L^2(X)$ with $(Tf)(x) = f(\tau(x))$. Thus τ is mixing.

[Hint: By linearity the hypothesis implies the conclusion whenever f and g are simple functions.]

21. Let $\mathbb{T}^d$ be the torus, and $\tau : x \mapsto x + \alpha$ the mapping arising in Exercise 17. Then τ is ergodic if and only if $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $\alpha_1, \alpha_2, \ldots, \alpha_d$, and 1 are linearly independent over the rationals. To do this show that:

(a) $\displaystyle\frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x)) \to \int_{\mathbb{T}^d} f(x)\,dx$ as $m \to \infty$, for each $x \in \mathbb{T}^d$, whenever f is continuous and periodic and α satisfies the hypothesis.

(b) Prove as a result that in this case τ is uniquely ergodic.

[Hint: Use (f) in Exercise 16.]

22. Let $X = \prod_{i=1}^{\infty} X_i$, where each (X_i, μ_i) is identical to (X_1, μ_1), with $\mu_1(X_1) = 1$, and let μ be the corresponding product measure defined in Exercise 15. Define the **shift** $\tau : X \to X$ by $\tau((x_1, x_2, \ldots)) = (x_2, x_3, \ldots)$ for $x = (x_i) \in \prod_{i=1}^{\infty} X_i$.

(a) Verify that τ is a measure-preserving transformation.

(b) Prove that τ is ergodic by showing that it is mixing.

(c) Note that in general τ is not uniquely ergodic.

If we define the corresponding shift on the two-sided infinite product, then τ is also a measure-preserving isomorphism.

[Hint: For (b) note that $\mu(\tau^{-n}(E \cap F)) = \mu(E)\mu(F)$ whenever E and F are cylinder sets and n is sufficiently large. For (c) note that, for example, if we fix a point $\overline{x} \in X_1$, the set $E = \{(x_i) : \; x_j = \overline{x} \text{ all } j\}$ is invariant.]

23. Let $X = \prod_{i=1}^{\infty} Z(2)$, where each factor is the two-point space $Z(2) = \{0, 1\}$ with $\mu_1(0) = \mu_1(1) = 1/2$, and suppose μ denotes the product measure on X. Consider the mapping $D : X \to [0, 1]$ given by $D(\{a_j\}) \to \sum_{j=1}^{\infty} \frac{a_j}{2^j}$. Then there are denumerable sets $Z_1 \subset X$ and $Z_2 \subset [0, 1]$, such that:

(a) D is a bijection from $X - Z_1$ to $[0, 1] - Z_2$.

(b) A set E in X is measurable if and only if $D(E)$ is measurable in $[0, 1]$, and $\mu(E) = m(D(E))$, where m is Lebesgue measure on $[0, 1]$.

(c) The shift map on $\prod_{i=1}^{\infty} Z(2)$ then becomes the doubling map of example (b) in Section 5.4.

24. Consider the following generalization of the doubling map. For each integer m, $m \geq 2$, we define the map τ_m of $(0, 1]$ by $\tau(x) = mx \mod 1$.

(a) Verify that τ is measure-preserving for Lebesgue measure.

(b) Show that τ is mixing, hence ergodic.

(c) Prove as a consequence that almost every number x is normal in the scale m, in the following sense. Consider the m-adic expansion of x,

$$x = \sum_{j=1}^{\infty} \frac{a_j}{m^j}, \qquad \text{where each } a_j \text{ is an integer } 0 \leq a_j \leq m - 1.$$

Then x is **normal** if for each integer k, $0 \leq k \leq m - 1$,

$$\frac{\#\{j : \; a_j = k, \; 1 \leq j \leq n\}}{N} \to \frac{1}{m} \qquad \text{as } N \to \infty.$$

Note the analogy with the equidistribution statements in Section 2, Chapter 4, of Book I.

25. Show that the mean ergodic theorem still holds if we replace the assumption that T is an isometry by the assumption that T is a **contraction**, that is, $\|Tf\| \leq \|f\|$ for all $f \in \mathcal{H}$.

[Hint: Prove that T is a contraction if and only if T^* is a contraction, and use the identity $(f, T^*f) = (Tf, f)$.]

26. There is an L^2 version of the maximal ergodic theorem. Suppose τ is a measure-preserving transformation on (X, μ). Here we do not assume that $\mu(X) < \infty$. Then

$$f^*(x) = \sup \frac{1}{m} \sum_{k=0}^{m-1} |f(\tau^k(x))|$$

satisfies

$$\|f^*\|_{L^2(X)} \leq c\|f\|_{L^2(X)}, \qquad \text{whenever } f \in L^2(X).$$

The proof is the same as outlined in Problem 6, Chapter 5 for the maximal function on $\mathbb{R}^d$. With this, extend the pointwise ergodic theorem to the case where $\mu(X) = \infty$, as follows:

(a) Show that $\lim_{m\to\infty} \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converges for a.e. x to $P(f)(x)$ for every $f \in L^2(X)$, because this holds for a dense subspace of $L^2(X)$.

(b) Prove that the conclusion holds for every $f \in L^1(X)$, because it holds for the dense subspace $L^1(X) \cap L^2(X)$.

27. We saw that if $\|f_n\|_{L^2} \leq 1$, then $\frac{f_n(x)}{n} \to 0$ as $n \to \infty$ for a.e. x. However, show that the analogue where one replaces the L^2-norm by the L^1-norm fails, by constructing a sequence $\{f_n\}$, $f_n \in L^1(X)$, $\|f_n\|_{L^1} \leq 1$, but with $\limsup_{n\to\infty} \frac{f_n(x)}{n} = \infty$ for a.e. x.

[Hint: Find intervals $I_n \subset [0,1]$, so that $m(I_n) = 1/(n \log n)$ but $\limsup_{n\to\infty}\{I_n\} = [0,1]$. Then take $f_n(x) = n \log n \chi_{I_n}$.]

28. We know by the Borel-Cantelli lemma that if $\{E_n\}$ is a collection of measurable sets in a measure a space (X, μ) and $\sum_{n=1}^{\infty} \mu(E_n) < \infty$ then $E = \limsup_{n\to\infty}\{E_n\}$ has measure zero.

In the opposite direction, if τ is a mixing measure-preserving transformation on X (with $\mu(X) = 1$), then whenever $\sum_{n=1}^{\infty} \mu(E_n) = \infty$, there are integers $m = m_n$ so that if $E_n' = \tau^{-m_n}(E_n)$, then $\limsup_{n\to\infty}(E_n') = X$, except for a set of measure 0.

8 Problems

1. Suppose Φ is a C^1 bijection of an open set $\mathcal{O}$ in $\mathbb{R}^d$ onto another open set $\mathcal{O}'$ in $\mathbb{R}^d$.

(a) If E is a measurable subset of $\mathcal{O}$, then $\Phi(E)$ is also measurable.

(b) $m(\Phi(E)) = \int_E |\det \Phi'(x)|\, dx$, where Φ' is the Jacobian of Φ.

(c) $\int_{\mathcal{O}'} f(y)\, dy = \int_{\mathcal{O}} f(\Phi(x))\, |\det \Phi'(x)|\, dx$ whenever f is integrable on $\mathcal{O}'$.

[Hint: To prove (a) follow the argument in Exercise 8, Chapter 1. For (b) assume E is a bounded open set, and write E as $\bigcup_{j=1}^{\infty} Q_j$, where Q_j are cubes whose interiors are disjoint, and whose diameters are less than ϵ. Let z_k be the center of Q_k. Then if $x \in Q_k$,

$$\Phi(x) = \Phi(z_k) + \Phi'(z_k)(x - z_k) + o(\epsilon),$$

hence $\Phi(Q_k) = \Phi(z_k) + \Phi'(z_k)(Q_k - z_k) + o(\epsilon)$, and as a result $(1 - \eta(\epsilon))\Phi'(z_k)(Q_k - z_k) \subset \Phi(Q_k) - \Phi(z_k) \subset (1 + \eta(\epsilon))\Phi'(z_k)(Q_k - z_k)$, where $\eta(\epsilon) \to 0$ as $\epsilon \to 0$. This means that

$$m(\Phi(\mathcal{O})) = \sum_k m(\Phi(Q_k)) = \sum_k |\det(\Phi'(z_k))|\, m(Q_k) + o(1) \qquad \text{as } \epsilon \to 0$$

on account of the linear transformation property of the Lebesgue measure given in Problem 4 of Chapter 2. Note that (b) is (c) for $f(\Phi(x)) = \chi_E(x)$.]

2. Show as a consequence of the previous problem: the measure $d\mu = \frac{dx\,dy}{y^2}$ in the upper half-plane $\mathbb{R}^2_+ = \{z = x + iy,\ y > 0\}$ is preserved by any fractional linear transformation $z \mapsto \frac{az+b}{cz+d}$, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ belongs to $\mathrm{SL}_2(\mathbb{R})$.

3. Let S be a hypersurface in $\mathbb{R}^d = \mathbb{R}^{d-1} \times \mathbb{R}$, given by

$$S = \{(x, y) \in \mathbb{R}^{d-1} \times \mathbb{R} : y = F(x)\},$$

with F a C^1 function defined on an open set Ω in $\mathbb{R}^{d-1}$. For each subset $E \subset \Omega$ we write $\widehat{E}$ for the corresponding subset of S given by $\widehat{E} = \{(x, F(x))\, x \in E\}$. We note that the Borel sets of S can be defined in terms of the metric on S (which is the restriction of the Euclidean metric on $\mathbb{R}^d$). Thus if E is a Borel set in Ω, then $\widehat{E}$ is a Borel subset of S.

(a) Let μ be the Borel measure on S given by

$$\mu(\widehat{E}) = \int_E \sqrt{1 + |\nabla F|^2}\, dx.$$

If B is a ball in Ω, let $\widehat{B}^\delta = \{(x, y) \in \mathbb{R}^d,\ d((x, y), \widehat{B}) < \delta\}$. Show that

$$\mu(\widehat{B}) = \lim_{\delta \to 0} \frac{1}{2\delta} m((\widehat{B})^\delta),$$

where m denotes the d-dimensional Lebesgue measure. This result is analogous to Theorem 4.4 in Chapter 3.

(b) One may apply (a) to the case when S is the (upper) half of the unit sphere in $\mathbb{R}^d$, given by $y = F(x)$, $F(x) = (1 - |x|^2)^{1/2}$, $|x| < 1$, $x \in \mathbb{R}^{d-1}$. Show that in this case $d\mu = d\sigma$, the measure on the sphere arising in the polar coordinate formula in Section 3.2.

(c) The above conclusion allows one to write an explicit formula for $d\sigma$ in terms of spherical coordinates. Take, for example, the case $d = 3$, and write $y = \cos\theta$, $x = (x_1, x_2) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi)$ with $0 \le \theta < \pi/2$, $0 \le \varphi < 2\pi$. Then according to (a) and (b) the element of area $d\sigma$ equals $(1 - |x|^2)^{-1/2}\, dx$. Use the change of variable theorem in Problem 1 to deduce that in this case $d\sigma = \sin\theta\, d\theta\, d\varphi$. This may be generalized to d dimensions, $d \ge 2$, to obtain the formulas in Section 2.4 of the appendix in Book I.

4.* Let μ be a Borel measure on the sphere S^{d-1} which is rotation-invariant in the following sense: $\mu(r(E)) = \mu(E)$, for every rotation r of $\mathbb{R}^d$ and each Borel subset E of S^{d-1}. If $\mu(S^{d-1}) < \infty$, then μ is a constant multiple of the measure σ arising in the polar coordinate integration formula.

[Hint: Show that

$$\int_{S^{d-1}} Y_k(x)\, d\mu(x) = 0$$

for every surface spherical harmonic of degree $k \geq 1$. As a result, there is a constant c so that

$$\int_{S^{d-1}} f\, d\mu = c \int_{S^{d-1}} f\, d\sigma$$

for every continuous function f on S^{d-1}.]

5.* Suppose X is a metric space, and μ is a Borel measure on X with the property that $\mu(B) < \infty$ for every ball B. Define $C_0(X)$ to be the vector space of continuous functions on X that are each supported in some closed ball. Then $\ell(f) = \int_X f\, d\mu$ defines a linear functional on $C_0(X)$ that is positive, that is, $\ell(f) \geq 0$ if $f \geq 0$.

Conversely, for any positive linear functional ℓ on $C_0(X)$, there exists a unique Borel measure μ that is finite on all balls, such that $\ell(f) = \int f\, d\mu$.

6. Consider an automorphism A of $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, that is, A is a linear isomorphism of $\mathbb{R}^d$ that preserves the lattice $\mathbb{Z}^d$. Note that A can be written as a $d \times d$ matrix whose entries are integers, with $\det A = \pm 1$. Define the mapping $\tau : \mathbb{T}^d \to \mathbb{T}^d$ by $\tau(x) = A(x)$.

(a) Observe that τ is a measure-preserving isomorphism of $\mathbb{T}^d$.

(b) Show that τ is ergodic (in fact, mixing) if and only if A has no eigenvalues of the form $e^{2\pi i p/q}$, where p and q are integers.

(c) Note that τ is never uniquely ergodic.

[Hint: The condition (b) is the same as $(A^t)^q$ has no invariant vectors, where A^t is the transpose of A. Note also that $f(\tau^k(x)) = e^{2\pi i (A^t)^k(n)\cdot x}$ where $f(x) = e^{2\pi i n\cdot x}$.]

7.* There is a version of the maximal ergodic theorem that is akin to the "rising sun lemma" and Exercise 6 in Chapter 3.

Suppose f is real-valued, and $f^{\#}(x) = \sup_m \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$. Let $E_0 = \{x : f^{\#}(x) > 0\}$. Then

$$\int_{E_0} f(x)\, dx \geq 0.$$

As a result (when we apply this to $f(x) - \alpha$), we get when $f \geq 0$ that

$$\mu\{x :\ f^*(x) > \alpha\} \leq \frac{1}{\alpha} \int_{\{f^*(x)>\alpha\}} f(x)\, dx.$$

In particular, the constant A in Theorem 5.3 can be taken to be 1.

8. Let $X = [0, 1)$, $\tau(x) = \langle 1/x \rangle$, $x \neq 0$, $\tau(0) = 0$. Here $\langle x \rangle$ denotes the fractional part of x. With the measure $d\mu = \frac{1}{\log 2} \frac{dx}{1+x}$, we have of course $\mu(X) = 1$.

Show that τ is a measure-preserving transformation.

[Hint: $\sum_{k=1}^{\infty} \frac{1}{(x+k)(x+k+1)} = \frac{1}{1+x}$.]

9.[*] The transformation τ in the previous problem is ergodic.

10.[*] The connection between continued fractions and the transformation $\tau(x) = \langle 1/x \rangle$ will now be described. A **continued fraction**, $a_0 + 1/(a_1 + 1/a_2) \cdots$, also written as $[a_0 a_1 a_2 \cdots]$, where the a_j are positive integers, can be assigned to any positive real number x in the following way. Starting with x, we successively transform it by two alternating operations: reducing it modulo 1 to lie in $[0, 1)$, and then taking the reciprocal of that number. The integers a_j that arise then define the continued fraction of x.

Thus we set $x = a_0 + r_0$, where $a_0 = [x] =$ the greatest integer in x, and $r_0 \in [0, 1)$. Next we write $1/r_0 = a_1 + r_1$, with $a_1 = [1/r_0]$, $r_1 \in [0, 1)$, to obtain successively $1/r_{n-1} = a_n + r_n$, where $a_n = [1/r_{n-1}]$, $r_n \in [0, 1)$. If $r_n = 0$ for some n, we write $a_k = 0$ for all $k > n$, and say that such a continued fraction **terminates**.

Note that if $0 \leq x < 1$, then $r_0 = x$ and $a_1 = [1/x]$, while $r_1 = \langle 1/x \rangle = \tau(x)$. More generally then, $a_k(x) = [1/\tau^{k-1}(x)] = a_1 \tau^{k-1}(x)$. The following properties of continued fractions of positive real numbers x are known:

(a) The continued fraction of x terminates if and only if x is rational.

(b) If $x = [a_0 a_1 \cdots a_n \cdots]$, and $x_N = [a_0 a_1 \cdots a_N 00 \cdots]$, then $x_N \to x$ as $N \to \infty$. The sequence $\{x_N\}$ gives essentially an optimal approximation of x by rationals.

(c) The continued fraction is periodic, that is, $a_{k+N} = a_k$ for some $N \geq 1$, and all sufficiently large k, if and only if x is an algebraic number of degree ≤ 2 over the rationals.

(d) One can conclude that $\frac{a_1 + a_2 + \cdots + a_n}{n} \to \infty$ as $n \to \infty$ for almost every x. In particular, the set of numbers x whose continued fractions $[a_0 a_1 \cdots a_n \cdots]$ are bounded has measure zero.

[Hint: For (d) apply a consequence of the pointwise ergodic theorem, which is as follows: Suppose $f \geq 0$, and $\int f \, d\mu = \infty$. If τ is ergodic, then $\frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x)) \to \infty$ for a.e. x as $m \to \infty$. In the present case take $f(x) = [1/x]$.]

7 Hausdorff Measure and Fractals

> Carathéodory developed a remarkably simple generalization of Lebesgue's measure theory which in particular allowed him to define the p-dimensional measure of a set in q-dimensional space. In what follows, I present a small addition.... a clarification of p-dimensional measure that leads immediately to an extension to non-integral p, and thus gives rise to sets of fractional dimension.
>
> *F. Hausdorff,* 1919

> I coined *fractal* from the Latin adjective *fractus.* The corresponding Latin verb *frangere* means to "break": to create irregular fragments.
>
> *B. Mandelbrot,* 1977

The deeper study of the geometric properties of sets often requires an analysis of their extent or "mass" that goes beyond what can be expressed in terms of Lebesgue measure. It is here that the notions of the dimension of a set (which can be fractional) and an associated measure play a crucial role.

Two initial ideas may help to provide an intuitive grasp of the concept of the dimension of a set. The first can be understood in terms of how the set replicates under scalings. Given the set E, let us suppose that for some positive number n we have that $nE = E_1 \cup \cdots \cup E_m$, where the sets E_j are m essentially disjoint congruent copies of E. Note that if E were a line segment this would hold with $m = n$; if E were a square, we would have $m = n^2$; if E were a cube, then $m = n^3$; etc. Thus, more generally, we might be tempted to say that E has dimension α if $m = n^\alpha$. Observe that if E is the Cantor set $\mathcal{C}$ in $[0, 1]$, then $3\mathcal{C}$ consists of 2 copies of $\mathcal{C}$, one in $[0, 1]$ and the other in $[2, 3]$. Here $n = 3$, $m = 2$, and we would be led to conclude that $\log 2/\log 3$ is the dimension of the Cantor set.

Another approach is relevant for curves that are not necessarily rectifiable. Start with a curve $\Gamma = \{\gamma(t) : a \leq t \leq b\}$, and for each $\epsilon > 0$ consider polygonal lines joining $\gamma(a)$ to $\gamma(b)$, whose vertices lie on successive points of Γ, with each segment not exceeding ϵ in length. Denote by $\#(\epsilon)$ the least number of segments that arise for such polygonal lines. If $\#(\epsilon) \approx \epsilon^{-1}$ as $\epsilon \to 0$, then Γ is rectifiable. However, $\#(\epsilon)$ may well grow more rapidly than ϵ^{-1} as $\epsilon \to 0$. If we had $\#(\epsilon) \approx \epsilon^{-\alpha}$, $1 < \alpha$, then, in the spirit of the previous example, it would be natural to say that Γ has dimension α. These considerations have even an interest in other parts of science. For instance, in studying the question of determining the length of the border of a country or its coastline, L.F. Richardson found that the length of the west coast of Britain obeyed the empirical law $\#(\epsilon) \approx \epsilon^{-\alpha}$, with α approximately 1.5. Thus one might conclude that the coast has fractional dimension!

While there are a number of different ways to make some of these heuristic notions precise, the theory that has the widest scope and greatest flexibility is the one involving Hausdorff measure and Hausdorff dimension. Probably the most elegant and simplest illustration of this theory can be seen in terms of its application to a general class of self-similar sets, and this is what we consider first. Among these are the curves of von Koch type, and these can have any dimension between 1 and 2.

Next, we turn to an example of a space-filling curve, which, broadly speaking, falls under the scope of self-replicating constructions. Not only does this curve have an intrinsic interest, but its nature reveals the important fact that from the point of view of measure theory the unit interval and the unit square are the same.

Our final topic is of a somewhat different nature. It begins with the realization of an unexpected regularity that all subsets of $\mathbb{R}^d$ (of finite Lebesgue measure) enjoy, when $d \geq 3$. This property fails in two dimensions, and the key counter-example is the Besicovitch set. This set appears also in a number of other problems. While it has measure zero, this is barely so, since its Hausdorff dimension is necessarily 2.

1 Hausdorff measure

The theory begins with the introduction of a new notion of volume or mass. This "measure" is closely tied with the idea of dimension which prevails throughout the subject. More precisely, following Hausdorff, one considers for each appropriate set E and each $\alpha > 0$ the quantity $m_\alpha(E)$, which can be interpreted as the α-dimensional mass of E among sets of dimension α, where the word "dimension" carries (for now) only

an intuitive meaning. Then, if α is larger than the dimension of the set E, the set has a negligible mass, and we have $m_\alpha(E) = 0$. If α is smaller than the dimension of E, then E is very large (comparatively), hence $m_\alpha(E) = \infty$. For the critical case when α is the dimension of E, the quantity $m_\alpha(E)$ describes the actual α-dimensional size of the set.

Two examples, to which we shall return in more detail later, illustrate this circle of ideas.

First, recall that the standard Cantor set $\mathcal{C}$ in $[0, 1]$ has zero Lebesgue measure. This statement expresses the fact that $\mathcal{C}$ has one-dimensional mass or length equal to zero. However, we shall prove that $\mathcal{C}$ has a well-defined fractional Hausdorff dimension of $\log 2/\log 3$, and that the corresponding Hausdorff measure of the Cantor set is positive and finite.

Another illustration of the theory developed below consists of starting with Γ, a rectifiable curve in the plane. Then Γ has zero two-dimensional Lebesgue measure. This is intuitively clear, since Γ is a one-dimensional object in a two-dimensional space. This is where the Hausdorff measure comes into play: the quantity $m_1(\Gamma)$ is not only finite, but precisely equal to the length of Γ as we defined it in Section 3.1 of Chapter 3.

We first consider the relevant exterior measure, defined in terms of coverings, whose restriction to the Borel sets is the desired Hausdorff measure.

For any subset E of $\mathbb{R}^d$, we define the **exterior α-dimensional Hausdorff measure** of E by

$$m_\alpha^*(E) = \lim_{\delta \to 0} \inf \left\{ \sum_k (\text{diam } F_k)^\alpha : E \subset \bigcup_{k=1}^\infty F_k, \quad \text{diam } F_k \leq \delta \text{ all } k \right\},$$

where diam S denotes the diameter of the set S, that is, diam $S = \sup\{|x - y| : x, y \in S\}$. In other words, for each $\delta > 0$ we consider covers of E by countable families of (arbitrary) sets with diameter less than δ, and take the infimum of the sum $\sum_k (\text{diam } F_k)^\alpha$. We then define $m_\alpha^*(E)$ as the limit of these infimums as δ tends to 0. We note that the quantity

$$\mathcal{H}_\alpha^\delta(E) = \inf \left\{ \sum_k (\text{diam } F_k)^\alpha : E \subset \bigcup_{k=1}^\infty F_k, \quad \text{diam } F_k \leq \delta \text{ all } k \right\}$$

is increasing as δ decreases, so that the limit

$$m_\alpha^*(E) = \lim_{\delta \to 0} \mathcal{H}_\alpha^\delta(E)$$

exists, although $m_\alpha^*(E)$ could be infinite. We note that in particular, one has $\mathcal{H}_\alpha^\delta(E) \leq m_\alpha^*(E)$ for all $\delta > 0$. When defining the exterior measure $m_\alpha^*(E)$ it is important to require that the coverings be of

sets of arbitrarily small diameters; this is the thrust of the definition $m_\alpha^*(E) = \lim_{\delta\to 0} \mathcal{H}_\alpha^\delta(E)$. This requirement, which is not relevant for Lebesgue measure, is needed to ensure the basic additive feature stated in Property 3 below. (See also Exercise 12.)

Scaling is the key notion that appears at the heart of the definition of the exterior Hausdorff measure. Loosely speaking, the measure of a set scales according to its dimension. For instance, if Γ is a one-dimensional subset of $\mathbb{R}^d$, say a smooth curve of length L, then $r\Gamma$ has total length rL. If Q is a cube in $\mathbb{R}^d$, the volume of rQ is $r^d|Q|$. This feature is captured in the definition of exterior Hausdorff measure by the fact that if the set F is scaled by r, then $(\text{diam } F)^\alpha$ scales by r^α. This key idea reappears in the study of self-similar sets in Section 2.2.

We begin with a list of properties satisfied by the Hausdorff exterior measure.

Property 1 (Monotonicity) *If $E_1 \subset E_2$, then $m_\alpha^*(E_1) \leq m_\alpha^*(E_2)$.*

This is straightforward, since any cover of E_2 is also a cover of E_1.

Property 2 (Sub-additivity) *$m_\alpha^*(\bigcup_{j=1}^\infty E_j) \leq \sum_{j=1}^\infty m_\alpha^*(E_j)$ for any countable family $\{E_j\}$ of sets in $\mathbb{R}^d$.*

For the proof, fix δ, and choose for each j a cover $\{F_{j,k}\}_{k=1}^\infty$ of E_j by sets of diameter less than δ such that $\sum_k (\text{diam } F_{j,k})^\alpha \leq \mathcal{H}_\alpha^\delta(E_j) + \epsilon/2^j$. Since $\bigcup_{j,k} F_{j,k}$ is a cover of E by sets of diameter less than δ, we must have

$$\begin{aligned}\mathcal{H}_\alpha^\delta(E) &\leq \sum_{j=1}^\infty \mathcal{H}_\alpha^\delta(E_j) + \epsilon \\ &\leq \sum_{j=1}^\infty m_\alpha^*(E_j) + \epsilon.\end{aligned}$$

Since ϵ is arbitrary, the inequality $\mathcal{H}_\alpha^\delta(E) \leq \sum m_\alpha^*(E_j)$ holds, and we let δ tend to 0 to prove the countable sub-additivity of m_α^*.

Property 3 *If $d(E_1, E_2) > 0$, then $m_\alpha^*(E_1 \cup E_2) = m_\alpha^*(E_1) + m_\alpha^*(E_2)$.*

It suffices to prove that $m_\alpha^*(E_1 \cup E_2) \geq m_\alpha^*(E_1) + m_\alpha^*(E_2)$ since the reverse inequality is guaranteed by sub-additivity. Fix $\epsilon > 0$ with $\epsilon < d(E_1, E_2)$. Given any cover of $E_1 \cup E_2$ with sets $F_1, F_2 \ldots$, of diameter less than δ, where $\delta < \epsilon$, we let

$$F_j' = E_1 \cap F_j \quad \text{and} \quad F_j'' = E_2 \cap F_j.$$

Then $\{F'_j\}$ and $\{F''_j\}$ are covers for E_1 and E_2, respectively, and are disjoint. Hence,

$$\sum_j (\text{diam } F'_j)^\alpha + \sum_i (\text{diam } F''_i)^\alpha \leq \sum_k (\text{diam } F_k)^\alpha.$$

Taking the infimum over the coverings, and then letting δ tend to zero yields the desired inequality.

At this point, we note that m^*_α satisfies all the properties of a metric Carathéodory exterior measure as discussed in Chapter 6. Thus m^*_α is a countably additive measure when restricted to the Borel sets. We shall therefore restrict ourselves to Borel sets and write $m_\alpha(E)$ instead of $m^*_\alpha(E)$. The measure m_α is called the **α-dimensional Hausdorff measure**.

Property 4 *If $\{E_j\}$ is a countable family of disjoint Borel sets, and $E = \bigcup_{j=1}^\infty E_j$, then*

$$m_\alpha(E) = \sum_{j=1}^\infty m_\alpha(E_j).$$

For what follows in this chapter, the full additivity in the above property is not needed, and we can manage with a weaker form whose proof is elementary and not dependent on the developments of Chapter 6. (See Exercise 2.)

Property 5 *Hausdorff measure is invariant under translations*

$$m_\alpha(E + h) = m_\alpha(E) \quad \textit{for all } h \in \mathbb{R}^d,$$

and rotations

$$m_\alpha(rE) = m_\alpha(E),$$

where r is a rotation in $\mathbb{R}^d$.

Moreover, it scales as follows:

$$m_\alpha(\lambda E) = \lambda^\alpha m_\alpha(E) \quad \textit{for all } \lambda > 0.$$

These conclusions follow once we observe that the diameter of a set S is invariant under translations and rotations, and satisfies $\text{diam}(\lambda S) = \lambda \,\text{diam}(S)$ for $\lambda > 0$.

We describe next a series of properties of Hausdorff measure, the first of which is immediate from the definitions.

Property 6 *The quantity $m_0(E)$ counts the number of points in E, while $m_1(E) = m(E)$ for all Borel sets $E \subset \mathbb{R}$. (Here m denotes the Lebesgue measure on $\mathbb{R}$.)*

In fact, note that in one dimension every set of diameter δ is contained in an interval of length δ (and for an interval its length equals its Lebesgue measure).

In general, d-dimensional Hausdorff measure in $\mathbb{R}^d$ is, up to a constant factor, equal to Lebesgue measure.

Property 7 *If E is a Borel subset of $\mathbb{R}^d$, then $c_d m_d(E) = m(E)$ for some constant c_d that depends only on the dimension d.*

The constant c_d equals $m(B)/(\text{diam } B)^d$, for the unit ball B; note that this ratio is the same for all balls B in $\mathbb{R}^d$, and so $c_d = v_d/2^d$ (where v_d denotes the volume of the unit ball). The proof of this property relies on the so-called iso-diametric inequality, which states that among all sets of a given diameter, the ball has largest volume. (See Problem 2.) Without using this geometric fact one can prove the following substitute.

Property 7′ *If E is a Borel subset of $\mathbb{R}^d$ and $m(E)$ is its Lebesgue measure, then $m_d(E) \approx m(E)$, in the sense that*

$$c_d m_d(E) \leq m(E) \leq 2^d c_d m_d(E).$$

Using Exercise 26 in Chapter 3 we can find for every $\epsilon, \delta > 0$, a covering of E by balls $\{B_j\}$, such that diam $B_j < \delta$, while $\sum_j m(B_j) \leq m(E) + \epsilon$. Now,

$$\mathcal{H}_d^\delta(E) \leq \sum_j (\text{diam } B_j)^d = c_d^{-1} \sum_j m(B_j) \leq c_d^{-1}(m(E) + \epsilon).$$

Letting δ and ϵ tend to 0, we get $m_d(E) \leq c_d^{-1} m(E)$. For the reverse direction, let $E \subset \bigcup_j F_j$ be a covering with $\sum_j (\text{diam } F_j)^d \leq m_d(E) + \epsilon$. We can always find closed balls B_j centered at a point of F_j so that $B_j \supset F_j$ and diam $B_j = 2 \,\text{diam } F_j$. However, $m(E) \leq \sum_j m(B_j)$, since $E \subset \bigcup_j B_j$, and the last sum equals

$$\sum c_d (\text{diam } B_j)^d = 2^d c_d \sum (\text{diam } F_j)^d \leq 2^d c_d \left(m_d(E) + \epsilon \right).$$

Letting $\epsilon \to 0$ gives $m(E) \leq 2^d c_d m_d(E)$.

Property 8 *If $m_\alpha^*(E) < \infty$ and $\beta > \alpha$, then $m_\beta^*(E) = 0$. Also, if $m_\alpha^*(E) > 0$ and $\beta < \alpha$, then $m_\beta^*(E) = \infty$.*

Indeed, if diam $F \leq \delta$, and $\beta > \alpha$, then

$$(\text{diam } F)^{\beta} = (\text{diam } F)^{\beta-\alpha}(\text{diam } F)^{\alpha} \leq \delta^{\beta-\alpha}(\text{diam } F)^{\alpha}.$$

Consequently

$$\mathcal{H}_{\beta}^{\delta}(E) \leq \delta^{\beta-\alpha}\mathcal{H}_{\alpha}^{\delta}(E) \leq \delta^{\beta-\alpha}m_{\alpha}^{*}(E).$$

Since $m_{\alpha}^{*}(E) < \infty$ and $\beta - \alpha > 0$, we find in the limit as δ tends to 0, that $m_{\beta}^{*}(E) = 0$.

The contrapositive gives $m_{\beta}^{*}(E) = \infty$ whenever $m_{\alpha}^{*}(E) > 0$ and $\beta < \alpha$.

We now make some easy observations that are consequences of the above properties.

1. If I is a finite line segment in $\mathbb{R}^d$, then $0 < m_1(I) < \infty$.
2. More generally, if Q is a k-cube in $\mathbb{R}^d$ (that is, Q is the product of k non-trivial intervals and $d-k$ points), then $0 < m_k(Q) < \infty$.
3. If $\mathcal{O}$ is a non-empty open set in $\mathbb{R}^d$, then $m_{\alpha}(\mathcal{O}) = \infty$ whenever $\alpha < d$. Indeed, this follows because $m_d(\mathcal{O}) > 0$.
4. Note that we can always take $\alpha \leq d$. This is because when $\alpha > d$, m_{α} vanishes on every ball, and hence on all of $\mathbb{R}^d$.

2 Hausdorff dimension

Given a Borel subset E of $\mathbb{R}^d$, we deduce from Property 8 that there exists a unique α such that

$$m_{\beta}(E) = \begin{cases} \infty & \text{if } \beta < \alpha, \\ 0 & \text{if } \alpha < \beta. \end{cases}$$

In other words, α is given by

$$\alpha = \sup\{\beta : m_{\beta}(E) = \infty\} = \inf\{\beta : m_{\beta}(E) = 0\}.$$

We say that E has **Hausdorff dimension** α, or more succinctly, that E has dimension α. We shall write $\alpha = \dim E$. At the critical value α we can say no more than that in general the quantity $m_{\alpha}(E)$ satisfies $0 \leq m_{\alpha}(E) \leq \infty$. If E is bounded and the inequalities are strict, that is, $0 < m_{\alpha}(E) < \infty$, we say that E has **strict Hausdorff dimension** α. The term **fractal** is commonly applied to sets of fractional dimension.

In general, calculating the Hausdorff measure of a set is a difficult problem. However, it is possible in some cases to bound this measure from above and below, and hence determine the dimension of the set in question. A few examples will illustrate these new concepts.

2.1 Examples

The Cantor set

The first striking example consists of the Cantor set $\mathcal{C}$, which was constructed in Chapter 1 by successively removing the middle-third intervals in $[0,1]$.

Theorem 2.1 *The Cantor set $\mathcal{C}$ has strict Hausdorff dimension $\alpha = \log 2/\log 3$.*

The inequality

$$m_\alpha(\mathcal{C}) \leq 1$$

follows from the construction of $\mathcal{C}$ and the definitions. Indeed, recall from Chapter 1 that $\mathcal{C} = \bigcap C_k$, where each C_k is a finite union of 2^k intervals of length 3^{-k}. Given $\delta > 0$, we first choose K so large that $3^{-K} < \delta$. Since the set C_K covers $\mathcal{C}$ and consists of 2^K intervals of diameter $3^{-K} < \delta$, we must have

$$\mathcal{H}_\alpha^\delta(\mathcal{C}) \leq 2^K (3^{-K})^\alpha.$$

However, α satisfies precisely $3^\alpha = 2$, hence $2^K(3^{-K})^\alpha = 1$, and therefore $m_\alpha(\mathcal{C}) \leq 1$.

The reverse inequality, which consists of proving that $0 < m_\alpha(\mathcal{C})$, requires a further idea. Here we rely on the Cantor-Lebesgue function, which maps $\mathcal{C}$ *surjectively* onto $[0,1]$. The key fact we shall use about this function is that it satisfies a precise continuity condition that reflects the dimension of the Cantor set.

A function f defined on a subset E of $\mathbb{R}^d$ satisfies a **Lipschitz condition** on E if there exists $M > 0$ such that

$$|f(x) - f(y)| \leq M|x - y| \quad \text{for all } x, y \in E.$$

More generally, a function f satisfies a **Lipschitz condition with exponent** γ (or is **Hölder** γ) if

$$|f(x) - f(y)| \leq M|x - y|^\gamma \quad \text{for all } x, y \in E.$$

The only interesting case is when $0 < \gamma \leq 1$. (See Exercise 3.)

Lemma 2.2 *Suppose a function f defined on a compact set E satisfies a Lipschitz condition with exponent γ. Then*

(i) $m_\beta(f(E)) \leq M^\beta m_\alpha(E)$ *if* $\beta = \alpha/\gamma$.

(ii) $\dim f(E) \leq \frac{1}{\gamma} \dim E$.

Proof. Suppose $\{F_k\}$ is a countable family of sets that covers E. Then $\{f(E \cap F_k)\}$ covers $f(E)$ and, moreover, $f(E \cap F_k)$ has diameter less than $M(\text{diam } F_k)^\gamma$. Hence

$$\sum_k (\text{diam } f(E \cap F_k))^{\alpha/\gamma} \leq M^{\alpha/\gamma} \sum_k (\text{diam } F_k)^\alpha,$$

and part (i) follows. This result now immediately implies conclusion (ii).

Lemma 2.3 *The Cantor-Lebesgue function F on $\mathcal{C}$ satisfies a Lipschitz condition with exponent* $\gamma = \log 2/\log 3$.

Proof. The function F was constructed in Section 3.1 of Chapter 3 as the limit of a sequence $\{F_n\}$ of piecewise linear functions. The function F_n increases by at most 2^{-n} on each interval of length 3^{-n}. So the slope of F_n is always bounded by $(3/2)^n$, and hence

$$|F_n(x) - F_n(y)| \leq \left(\frac{3}{2}\right)^n |x - y|.$$

Moreover, the approximating sequence also satisfies $|F(x) - F_n(x)| \leq 1/2^n$. These two estimates together with an application of the triangle inequality give

$$\begin{aligned} |F(x) - F(y)| &\leq |F_n(x) - F_n(y)| + |F(x) - F_n(x)| + |F(y) - F_n(y)| \\ &\leq \left(\frac{3}{2}\right)^n |x - y| + \frac{2}{2^n}. \end{aligned}$$

Having fixed x and y, we then minimize the right hand side by choosing n so that both terms have the same order of magnitude. This is achieved by taking n so that $3^n|x - y|$ is between 1 and 3. Then, we see that

$$|F(x) - F(y)| \leq c2^{-n} = c(3^{-n})^\gamma \leq M|x - y|^\gamma,$$

since $3^\gamma = 2$ and 3^{-n} is not greater than $|x - y|$. This argument is repeated in Lemma 2.8 below.

With $E = \mathcal{C}$, f the Cantor-Lebesgue function, and $\alpha = \gamma = \log 2/\log 3$, the two lemmas give

$$m_1([0, 1]) \leq M^\beta m_\alpha(\mathcal{C}).$$

Thus $m_\alpha(\mathcal{C}) > 0$, and we find that $\dim \mathcal{C} = \log 2/\log 3$.

The proof of this example is typical in the sense that the inequality $m_\alpha(\mathcal{C}) < \infty$ is usually easier to obtain than $0 < m_\alpha(\mathcal{C})$. Also, with some extra effort, it is possible to show that the $\log 2/\log 3$-dimensional Hausdorff measure of $\mathcal{C}$ is precisely 1. (See Exercise 7.)

Rectifiable curves

A further example of the role of dimension comes from looking at continuous curves in $\mathbb{R}^d$. Recall that a continuous curve $\gamma : [a, b] \to \mathbb{R}^d$ is said to be **simple** if $\gamma(t_1) \neq \gamma(t_2)$ whenever $t_1 \neq t_2$, and **quasi-simple** if the mapping $t \mapsto z(t)$ is injective for t in the complement of finitely many points.

Theorem 2.4 *Suppose the curve γ is continuous and quasi-simple. Then γ is rectifiable if and only if $\Gamma = \{\gamma(t) : a \leq t \leq b\}$ has strict Hausdorff dimension one. Moreover, in this case the length of the curve is precisely its one-dimensional measure $m_1(\Gamma)$.*

Proof. Suppose to begin with that Γ is a rectifiable curve of length L, and consider an arc-length parametrization $\tilde{\gamma}$ such that $\Gamma = \{\tilde{\gamma}(t) : 0 \leq t \leq L\}$. This parametrization satisfies the Lipschitz condition

$$|\tilde{\gamma}(t_1) - \tilde{\gamma}(t_2)| \leq |t_1 - t_2|.$$

This follows since $|t_1 - t_2|$ is the length of the curve between t_1 and t_2, which is greater than the distance from $\tilde{\gamma}(t_1)$ to $\tilde{\gamma}(t_2)$. Since $\tilde{\gamma}$ satisfies the conditions of Lemma 2.2 with exponent 1 and $M = 1$, we find that

$$m_1(\Gamma) \leq L.$$

To prove the reverse inequality, we let $a = t_0 < t_1 < \cdots < t_N = b$ denote a partition of $[a, b]$ and let

$$\Gamma_j = \{\gamma(t) : t_j \leq t \leq t_{j+1}\},$$

so that $\Gamma = \bigcup_{j=0}^{N-1} \Gamma_j$, and hence

$$m_1(\Gamma) = \sum_{j=0}^{N-1} m_1(\Gamma_j)$$

by an application of Property 4 of the Hausdorff measure and the fact that Γ is quasi-simple. Indeed, by removing finitely many points the

union $\bigcup_{j=0}^{N-1} \Gamma_j$ becomes disjoint, while the points removed clearly have zero m_1-measure. We next claim that $m_1(\Gamma_j) \geq \ell_j$, where ℓ_j is the distance from $\gamma(t_j)$ to $\gamma(t_{j+1})$, that is, $\ell_j = |\gamma(t_{j+1}) - \gamma(t_j)|$. To see this, recall that Hausdorff measure is rotation-invariant, and introduce new orthogonal coordinates x and y such that $[\gamma(t_j), \gamma(t_j + 1)]$ is the segment $[0, \ell_j]$ on the x-axis. The projection $\pi(x, y) = x$ satisfies the Lipschitz condition

$$|\pi(P) - \pi(Q)| \leq |P - Q|,$$

and clearly the segment $[0, \ell_j]$ on the x-axis is contained in the image $\pi(\Gamma_j)$. Therefore, Lemma 2.2 guarantees

$$\ell_j \leq m_1(\Gamma_j),$$

and thus $m_1(\Gamma) \geq \sum \ell_j$. Since by definition the length L of Γ is the supremum of the sums $\sum \ell_j$ over all partitions of $[a, b]$, we find that $m_1(\Gamma) \geq L$, as desired.

Conversely, if Γ has strict Hausdorff dimension 1, then $m_1(\Gamma) < \infty$, and the above argument shows that Γ is rectifiable.

The reader may note the resemblance of this characterization of rectifiability and an earlier one in terms of Minkowski content, given in Chapter 3. In this connection we point out that there is a different notion of dimension that is sometimes used instead of Hausdorff dimension. For a compact set E, this dimension is given in terms of the size of $E^\delta = \{x \in \mathbb{R}^d : \ d(x, E) < \delta\}$ as $\delta \to 0$. One observes that if E is a k-dimensional cube in $\mathbb{R}^d$, then $m(E^\delta) \leq c\delta^{d-k}$ as $\delta \to 0$, with m the Lebesgue measure of $\mathbb{R}^d$. With this in mind, the **Minkowski dimension** of E is defined by

$$\inf \{\beta : \ m(E^\delta) = O(\delta^{d-\beta}) \text{ as } \delta \to 0\}.$$

One can show that the Hausdorff dimension of a set does not exceed its Minkowski dimension, but that equality does not hold in general. More details may be found in Exercises 17 and 18.

The Sierpinski triangle

A Cantor-like set can be constructed in the plane as follows. We begin with a (solid) closed equilateral triangle S_0, whose sides have unit length. Then, as a first step we remove the shaded open equilateral triangle pictured in Figure 1.

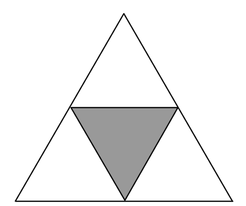
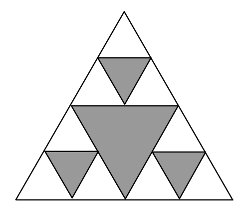
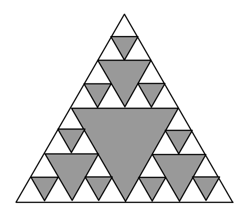
  

Figure 1. Construction of the Sierpinski triangle

This leaves three closed triangles whose union we denote by S_1. Each triangle is half the size of the original (or parent) triangle S_0, and these smaller closed triangles are said to be of the first **generation**: the triangles in S_1 are the children of the parent S_0. In the second step, we repeat the process in each triangle of the first generation. Each such triangle has three children of the second generation. We denote by S_2 the union of the three triangles in the second generation. We then repeat this process to find a sequence S_k of compact sets which satisfy the following properties:

(a) Each S_k is a union of 3^k closed equilateral triangles of side length 2^{-k}. (These are the triangles of the kth generation.)

(b) $\{S_k\}$ is a decreasing sequence of compact sets; that is, $S_{k+1} \subset S_k$ for all $k \geq 0$.

The **Sierpinski triangle** is the compact set defined by

$$\mathcal{S} = \bigcap_{k=0}^{\infty} S_k.$$

Theorem 2.5 *The Sierpinski triangle $\mathcal{S}$ has strict Hausdorff dimension $\alpha = \log 3/\log 2$.*

The inequality $m_\alpha(\mathcal{S}) \leq 1$ follows immediately from the construction. Given $\delta > 0$, choose K so that $2^{-K} < \delta$. Since the set S_K covers $\mathcal{S}$ and consists of 3^K triangles each of diameter $2^{-K} < \delta$, we must have

$$\mathcal{H}_\alpha^\delta(\mathcal{S}) \leq 3^K (2^{-K})^\alpha.$$

But since $2^\alpha = 3$, we find $\mathcal{H}_\alpha^\delta(\mathcal{S}) \leq 1$, hence $m_\alpha(\mathcal{S}) \leq 1$.

The inequality $m_\alpha(\mathcal{S}) > 0$ is more subtle. For its proof we need to fix a special point in each triangle that appears in the construction of $\mathcal{S}$.

We choose to call the lower left vertex of a triangle *the* vertex of that triangle. With this choice there are 3^k vertices of the kth generation. The argument that follows is based on the important fact that all these vertices belong to $\mathcal{S}$.

Suppose $\mathcal{S} \subset \bigcup_{j=1}^{\infty} F_j$, with diam $F_j < \delta$. We wish to prove that

$$\sum_j (\text{diam } F_j)^\alpha \geq c > 0$$

for some constant c. Clearly, each F_j is contained in a ball of twice the diameter of F_j, so upon replacing 2δ by δ and noting that $\mathcal{S}$ is compact, it suffices to show that if $\mathcal{S} \subset \bigcup_{j=1}^{N} B_j$, where $\mathcal{B} = \{B_j\}_{j=1}^{N}$ is a finite collection of balls whose diameters are less than δ, then

$$\sum_{j=1}^{N} (\text{diam } B_j)^\alpha \geq c > 0.$$

Suppose we have such a covering by balls. Consider the minimum diameter of the B_j, and choose k so that

$$2^{-k} \leq \min_{1 \leq j \leq N} \text{diam } B_j < 2^{-k+1}.$$

Lemma 2.6 *Suppose B is a ball in the covering $\mathcal{B}$ that satisfies*

$$2^{-\ell} \leq \text{diam } B < 2^{-\ell+1} \quad \textit{for some } \ell \leq k.$$

Then B contains at most $c3^{k-\ell}$ vertices of the kth generation.

In this chapter, we shall continue use the common practice of denoting by $c, c', \ldots$ generic constants whose values are unimportant and may change from one usage to another. We also use $A \approx B$ to denote that the quantities A and B are **comparable**, that is, $cB \leq A \leq c'B$, for appropriate constants c and c'.

Proof of Lemma 2.6. Let B^* denote the ball with same center as B but three times its diameter, and let $\triangle_k$ be a triangle of the kth generation whose vertex v lies in B. If $\triangle'_\ell$ denotes the triangle of the ℓth generation that contains $\triangle_k$, then since diam $B \geq 2^{-\ell}$,

$$v \in \triangle_k \subset \triangle'_\ell \subset B^*,$$

as shown in Figure 2.

Next, there is a positive constant c such that B^* can contain at most c distinct triangles of the ℓth generation. This is because triangles of the

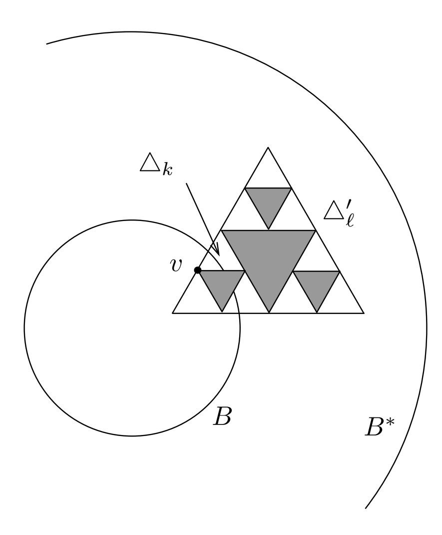


Figure 2. The setting in Lemma 2.6

ℓ^{th} generation have disjoint interiors and area equal to $c'4^{-\ell}$, while B^* has area at most equal to $c''4^{-\ell}$. Finally, each $\triangle'_\ell$ contains $3^{k-\ell}$ triangles of the kth generation, hence B can contain at most $c3^{k-\ell}$ vertices of triangles of the kth generation.

To complete the proof that $\sum_{j=1}^N (\text{diam } B_j)^\alpha \geq c > 0$, note that

$$\sum_{j=1}^N (\text{diam } B_j)^\alpha \geq \sum_\ell N_\ell 2^{-\ell\alpha},$$

where N_ℓ denotes the number of balls in $\mathcal{B}$ that satisfy $2^{-\ell} \leq \text{diam } B_j \leq 2^{-\ell+1}$. By the lemma, we see that the total number of vertices of triangles in the kth generation that can be covered by the collection $\mathcal{B}$ can be no more than $c\sum_\ell N_\ell 3^{k-\ell}$. Since all 3^k vertices of triangles in the kth generation belong to $\mathcal{S}$, and all vertices of the kth generation must be covered, we must have $c\sum_\ell N_\ell 3^{k-\ell} \geq 3^k$. Hence

$$\sum_\ell N_\ell 3^{-\ell} \geq c.$$

It now suffices to recall the definition of α which guarantees $2^{-\ell\alpha} = 3^{-\ell}$, and therefore

$$\sum_{j=1}^N (\text{diam } B_j)^\alpha \geq c,$$

as desired.

We give a final example that exhibits properties similar to the Cantor set and Sierpinski triangle. It is the curve discovered by von Koch in 1904.

The von Koch curve

Consider the unit interval $K_0 = [0, 1]$, which we may think of as lying on the x-axis in the xy-plane. Then consider the polygonal path K_1 illustrated in Figure 3, which consists of four equal line segments of length 1/3.

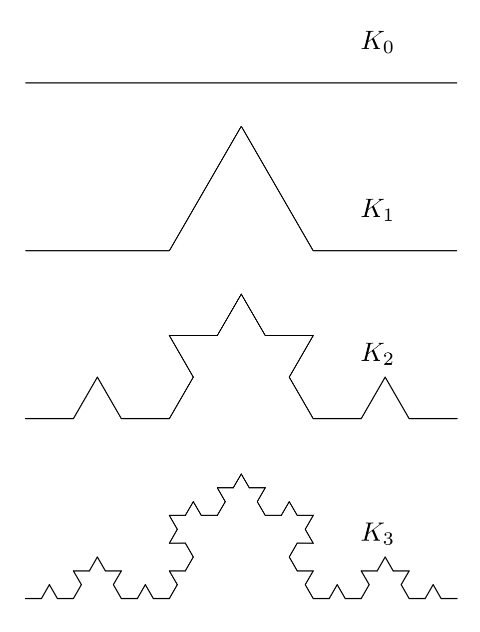


Figure 3. The first few stages in the construction of the von Koch curve

Let $K_1(t)$, for $0 \leq t \leq 1$, denote the parametrization of K_1 that has constant speed. In other words, as t travels from 0 to 1/4, the point $K_1(t)$ travels on the first line segment. As t travels from 1/4 to 1/2, the point $K_1(t)$ travels on the second line segment, and so on. In particular, we see that $K_1(\ell/4)$ for $0 \leq \ell \leq 4$ correspond to the five vertices of K_1.

At the second stage of the construction we repeat the process of replacing each line segment in stage one by the corresponding polygonal line. We then obtain the polygonal curve K_2 illustrated in Figure 3. It has $16 = 4^2$ segments of length $1/9 = 3^{-2}$. We choose a parametrization

$K_2(t)$ $(0 \le t \le 1)$ of K_2 that has constant speed. Observe that $K_2(\ell/4^2)$ for $0 \le \ell \le 4^2$ gives all vertices of K_2, and that the vertices of K_1 belong to K_2, with

$$K_2(\ell/4) = K_1(\ell/4) \quad \text{for } 0 \le \ell \le 4.$$

Repeating this process indefinitely, we obtain a sequence of continuous polygonal curves $\{K_j\}$, where K_j consists of 4^j segments of length 3^{-j} each. If $K_j(t)$ $(0 \le t \le 1)$ is the parametrization of K_j that has constant speed, then the vertices are precisely at the points $K_j(\ell/4^j)$, and

$$K_{j'}(\ell/4^j) = K_j(\ell/4^j) \quad \text{for } 0 \le \ell \le 4^j$$

whenever $j' \ge j$.

In the limit as j tends to infinity, the polygonal lines K_j tend to the **von Koch curve** $\mathcal{K}$. Indeed, we have

$$|K_{j+1}(t) - K_j(t)| \le 3^{-j} \quad \text{for all } 0 \le t \le 1 \text{ and } j \ge 0.$$

This is clear when $j = 0$, and follows by induction in j when we consider the nature of the construction of the jth stage. Since we may write

$$K_J(t) = K_1(t) + \sum_{j=1}^{J-1}(K_{j+1}(t) - K_j(t)),$$

the above estimate proves that the series

$$K_1(t) + \sum_{j=1}^{\infty}(K_{j+1}(t) - K_j(t))$$

converges absolutely and uniformly to a continuous function $\mathcal{K}(t)$ that is a parametrization of $\mathcal{K}$. Besides continuity, the function $\mathcal{K}(t)$ satisfies a regularity assumption that takes the form of a Lipschitz condition, as in the case of the Cantor-Lebesgue function.

Theorem 2.7 *The function $\mathcal{K}(t)$ satisfies a Lipschitz condition of exponent $\gamma = \log 3/\log 4$, that is:*

$$|\mathcal{K}(t) - \mathcal{K}(s)| \le M|t-s|^{\gamma} \quad \textit{for all } t, s \in [0,1].$$

We have already observed that $|K_{j+1}(t) - K_j(t)| \le 3^{-j}$. Since K_j travels a distance of 3^{-j} in 4^{-j} units of time, we see that

$$|K_j'(t)| \le \left(\frac{4}{3}\right)^j \quad \text{except when } t = \ell/4^j.$$

Consequently we must have

$$|K_j(t) - K_j(s)| \leq \left(\frac{4}{3}\right)^j |t - s|.$$

Moreover, $\mathcal{K}(t) = K_1(t) + \sum_{j=1}^{\infty}(K_{j+1}(t) - K_j(t))$. We now find ourselves in precisely the same situation as in the proof that the Cantor-Lebesgue function satisfies a Lipschitz condition with exponent $\log 2/\log 3$. We generalize that argument in the following lemma.

Lemma 2.8 *Suppose* $\{f_j\}$ *is a sequence of continuous functions on the interval* $[0, 1]$ *that satisfy*

$$|f_j(t) - f_j(s)| \leq A^j |t - s| \quad \textit{for some } A > 1,$$

and

$$|f_j(t) - f_{j+1}(t)| \leq B^{-j} \quad \textit{for some } B > 1.$$

Then the limit $f(t) = \lim_{j \to \infty} f_j(t)$ *exists and satisfies*

$$|f(t) - f(s)| \leq M|t - s|^{\gamma},$$

where $\gamma = \log B/\log(AB)$.

Proof. The continuous limit f is given by the uniformly convergent series

$$f(t) = f_1(t) + \sum_{k=1}^{\infty}(f_{k+1}(t) - f_k(t)),$$

and therefore

$$|f(t) - f_j(t)| \leq \sum_{k=j}^{\infty} |f_{k+1}(t) - f_k(t)| \leq \sum_{k=j}^{\infty} B^{-k} \leq cB^{-j}.$$

The triangle inequality, an application of the inequality just obtained, and the inequality in the statement of the lemma give

$$\begin{aligned} |f(t) - f(s)| &\leq |f_j(t) - f_j(s)| + |(f - f_j)(t)| + |(f - f_j)(s)| \\ &\leq c(A^j|t - s| + B^{-j}). \end{aligned}$$

For a fixed pair of numbers t and s with $t \neq s$, we choose j to minimize the sum $A^j|t - s| + B^{-j}$. This is essentially achieved by picking j so that

two terms $A^j|t-s|$ and B^{-j} are comparable. More precisely, we choose a j that satisfies

$$(AB)^j|t-s| \leq 1 \quad \text{and} \quad 1 \leq (AB)^{j+1}|t-s|.$$

Since $|t-s| \leq 2$ and $AB > 1$, such a j must exist. The first inequality then gives

$$A^j|t-s| \leq B^{-j},$$

while raising the second inequality to the power γ, and using the fact that $(AB)^\gamma = B$ gives

$$1 \leq B^j|t-s|^\gamma.$$

Thus $B^{-j} \leq |t-s|^\gamma$, and consequently

$$|f(t)-f(s)| \leq c(A^j|t-s| + B^{-j}) \leq M|t-s|^\gamma,$$

as was to be shown.

In particular, this result with Lemma 2.2 implies that

$$\dim \mathcal{K} \leq \frac{1}{\gamma} = \frac{\log 4}{\log 3}.$$

To prove that $m_\gamma(\mathcal{K}) > 0$ and hence $\dim \mathcal{K} = \log 4/\log 3$ requires an argument similar to the one given for the Sierpinski triangle. In fact, this argument generalizes to cover a general family of sets that have a self-similarity property. We therefore turn our attention to this general theory next.

Remarks. We mention some further facts about the von Koch curve. More details can be found in Exercises 13, 14, and 15 below.

1. The curve $\mathcal{K}$ is one in a family of similarly constructed curves. For each ℓ, $1/4 < \ell < 1/2$, consider at the first stage the curve $K_1^\ell(t)$ given by four line segments each of length ℓ, the first and last on the x-axis, and the second and third forming two sides of an isoceles triangle whose base lies on the x-axis. (See Figure 4.) The case $\ell = 1/3$ corresponds to the previously defined von Koch curve.

   Proceeding as in the case $\ell = 1/3$, one obtains a curve $\mathcal{K}^\ell$, and it can be seen that

$$\dim(\mathcal{K}^\ell) = \frac{\log 4}{\log 1/\ell}.$$

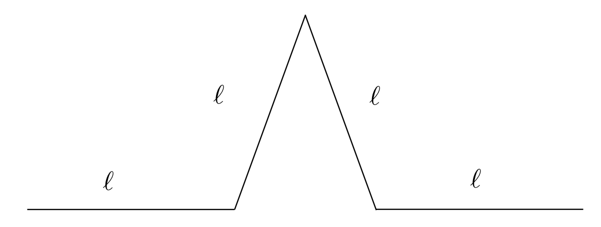


Figure 4. The curve $K_1^\ell(t)$

Thus for every α, $1 < \alpha < 2$, we have a curve of this kind of dimension α. Note that when $\ell \to 1/4$ the limiting curve is a straight line segment, which has dimension 1. When $\ell \to 1/2$, the limit can be seen to correspond to a "space-filling" curve.

2. The curves $t \mapsto \mathcal{K}^\ell(t)$, $1/4 < \ell \leq 1/2$, are each nowhere differentiable. One can also show that each curve is simple when $1/4 \leq \ell < 1/2$.

2.2 Self-similarity

The Cantor set $\mathcal{C}$, the Sierpinski triangle $\mathcal{S}$, and von Koch curve $\mathcal{K}$ all share an important property: each of these sets contains scaled copies of itself. Moreover, each of these examples was constructed by iterating a process closely tied to its scaling. For instance, the interval $[0, 1/3]$ contains a copy of the Cantor set scaled by a factor of 1/3. The same is true for the interval $[2/3, 1]$, and therefore

$$\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2,$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are scaled versions of $\mathcal{C}$. Also, each interval $[0, 1/9]$, $[2/9, 3/9]$, $[6/9, 7/9]$ and $[8/9, 1]$ contains a copy of $\mathcal{C}$ scaled by a factor of 1/9, and so on.

In the case of the Sierpinski triangle, each of the three triangles in the first generation contains a copy of $\mathcal{S}$ scaled by the factor of 1/2. Hence

$$\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3,$$

where each $\mathcal{S}_j$, $j = 1, 2, 3$, is obtained by scaling and translating the original Sierpinski triangle. More generally, every triangle in the kth generation is a copy of $\mathcal{S}$ scaled by the factor of $1/2^k$.

Finally, each line segment in the initial stage of the construction of the von Koch curve gives rise to a scaled and possibly rotated copy of the

von Koch curve. In fact

$$\mathcal{K} = \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3 \cup \mathcal{K}_4,$$

where $\mathcal{K}_j$, $j = 1, 2, 3, 4$, is obtained by scaling $\mathcal{K}$ by the factor of $1/3$ and translating and rotating it.

Thus these examples each contain replicas of themselves, but on a smaller scale. In this section, we give a precise definition of the resulting notion of self-similarity and prove a theorem determining the Hausdorff dimension of these sets.

A mapping $S : \mathbb{R}^d \to \mathbb{R}^d$ is said to be a **similarity** with **ratio** $r > 0$ if

$$|S(x) - S(y)| = r|x - y|.$$

It can be shown that every similarity of $\mathbb{R}^d$ is the composition of a translation, a rotation, and a dilation by r. (See Problem 3.)

Given finitely many similarities $S_1, \ldots, S_m$ with the same ratio r, we say that the set $F \subset \mathbb{R}^d$ is **self-similar** if

$$F = S_1(F) \cup \cdots \cup S_m(F).$$

We point out the relevance of the various examples we have already seen.

When $F = \mathcal{C}$ is the Cantor set, there are two similarities given by

$$S_1(x) = x/3 \quad \text{and} \quad S_2(x) = x/3 + 2/3$$

of ratio $1/3$. So $m = 2$ and $r = 1/3$.

In the case of $F = \mathcal{S}$, the Sierpinski triangle, the ratio is $r = 1/2$ and there are $m = 3$ similarities given by

$$S_1(x) = \frac{x}{2}, \quad S_2(x) = \frac{x}{2} + \alpha \quad \text{and} \quad S_3(x) = \frac{x}{2} + \beta.$$

Here, α and β are the points drawn in the first diagram in Figure 5.

If $F = \mathcal{K}$, the von Koch curve, we have

$$S_1(x) = \frac{x}{3}, \quad S_2(x) = \rho\frac{x}{3} + \alpha, \quad S_3(x) = \rho^{-1}\frac{x}{3} + \beta,$$

and

$$S_4(x) = \frac{x}{3} + \gamma,$$

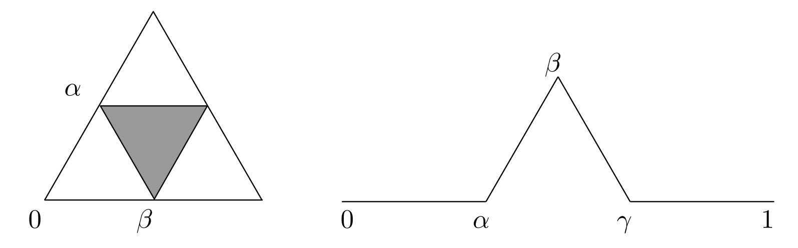


Figure 5. Similarities of the Sierpinski triangle and von Koch curve

where ρ is the rotation centered at the origin and of angle $\pi/3$. There are $m = 4$ similarities which have ratio $r = 1/3$. The points α, β, and γ are shown in the second diagram in Figure 5.

Another example, sometimes called the **Cantor dust** $\mathcal{D}$, is another two-dimensional version of the standard Cantor set. For each fixed $0 < \mu < 1/2$, the set $\mathcal{D}$ may be constructed by starting with the unit square $Q = [0,1] \times [0,1]$. At the first stage we remove everything but the four open squares in the corners of Q that have side length μ. This yields a union D_1 of four squares, as illustrated in Figure 6.

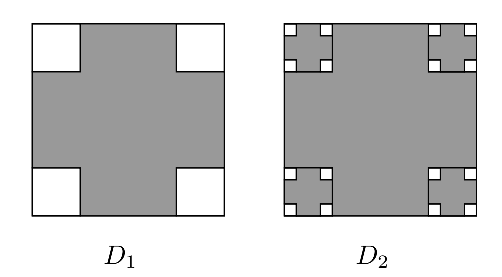


Figure 6. Construction of the Cantor dust

We repeat this process in each sub-square of D_1; that is, we remove everything but the four squares in the corner, each of side length μ^2. This gives a union D_2 of 16 squares. Repeating this process, we obtain a family $D_1 \supset D_2 \supset \cdots \supset D_k \supset \cdots$ of compact sets whose intersection defines the Cantor dust corresponding to the parameter μ.

There are here $m = 4$ similarities of ratio μ given by

$$\begin{aligned} S_1(x) &= \mu x, \\ S_2(x) &= \mu x + (0, 1-\mu), \\ S_3(x) &= \mu x + (1-\mu, 1-\mu), \\ S_4(x) &= \mu x + (1-\mu, 0). \end{aligned}$$

It is to be noted that $\mathcal{D}$ is the product $\mathcal{C}_\xi \times \mathcal{C}_\xi$, with $\mathcal{C}_\xi$ the Cantor set of constant dissection ξ, as defined in Exercise 3, of Chapter 1. Here $\xi = 1 - 2\mu$.

The first result we prove guarantees the existence of self-similar sets under the assumption that the similarities are contracting, that is, that their ratio satisfies $r < 1$.

Theorem 2.9 *Suppose $S_1, S_2, \ldots, S_m$ are m similartities, each with the same ratio r that satisfies $0 < r < 1$. Then there exists a unique non-empty compact set F such that*

$$F = S_1(F) \cup \cdots \cup S_m(F).$$

The proof of this theorem is in the nature of a fixed point argument. We shall begin with some large ball B and iteratively apply the mappings $S_1, \ldots, S_m$. The fact that the similarities have ratio $r < 1$ will suffice to imply that this process contracts to a unique set F with the desired property.

Lemma 2.10 *There exists a closed ball B so that $S_j(B) \subset B$ for all $j = 1, \ldots, m$.*

Proof. Indeed, we note that if S is a similarity with ratio r, then

$$\begin{aligned} |S(x)| &\leq |S(x) - S(0)| + |S(0)| \\ &\leq r|x| + |S(0)|. \end{aligned}$$

If we require that $|x| \leq R$ implies $|S(x)| \leq R$, it suffices to choose R so that $rR + |S(0)| \leq R$, that is, $R \geq |S(0)|/(1-r)$. In this fashion, we obtain for each S_j a ball B_j centered at the origin that satisfies $S_j(B_j) \subset B_j$. If B denotes the ball among the B_j with the largest radius, then the above shows that $S_j(B) \subset B$ for all j.

Now for any set A, let $\tilde{S}(A)$ denote the set given by

$$\tilde{S}(A) = S_1(A) \cup \cdots \cup S_m(A).$$

Note that if $A \subset A'$, then $\tilde{S}(A) \subset \tilde{S}(A')$.

Also observe that while each S_j is a mapping from $\mathbb{R}^d$ to $\mathbb{R}^d$, the mapping $\tilde{S}$ is not a point mapping, but takes subsets of $\mathbb{R}^d$ to subsets of $\mathbb{R}^d$.

To exploit the notion of contraction with a ratio less than 1, we introduce the distance between two compact sets as follows. For each $\delta > 0$ and set A, we let

$$A^\delta = \{x : d(x, A) < \delta\}.$$

Hence A^δ is a set that contains A but which is slightly larger in terms of δ. If A and B are two compact sets, we define the **Hausdorff distance** as

$$\operatorname{dist}(A, B) = \inf\{\delta : \ B \subset A^\delta \text{ and } A \subset B^\delta\}.$$

Lemma 2.11 *The distance function* dist *defined on compact subsets of* $\mathbb{R}^d$ *satisfies*

(i) $\operatorname{dist}(A, B) = 0$ *if and only if* $A = B$.

(ii) $\operatorname{dist}(A, B) = \operatorname{dist}(B, A)$.

(iii) $\operatorname{dist}(A, B) \leq \operatorname{dist}(A, C) + \operatorname{dist}(C, B)$.

If $S_1, \ldots, S_m$ *are similarities with ratio* r*, then*

(iv) $\operatorname{dist}(\tilde{S}(A), \tilde{S}(B)) \leq r \operatorname{dist}(A, B)$.

The proof of the lemma is simple and may be left to the reader.

Using both lemmas we may now prove Theorem 2.9. We first choose B as in Lemma 2.10, and let $F_k = \tilde{S}^k(B)$, where $\tilde{S}^k$ denotes the kth composition of $\tilde{S}$, that is, $\tilde{S}^k = \tilde{S}^{k-1} \circ \tilde{S}$ with $\tilde{S}^1 = \tilde{S}$. Each F_k is compact, non-empty, and $F_k \subset F_{k-1}$, since $\tilde{S}(B) \subset B$. If we let

$$F = \bigcap_{k=1}^{\infty} F_k,$$

then F is compact, non-empty, and clearly $\tilde{S}(F) = F$, since applying $\tilde{S}$ to $\bigcap_{k=1}^{\infty} F_k$ yields $\bigcap_{k=2}^{\infty} F_k$, which also equals F.

Uniqueness of the set F is proved as follows. Suppose G is another compact set so that $\tilde{S}(G) = G$. Then, an application of part (iv) in Lemma 2.11 yields $\operatorname{dist}(F, G) \leq r \operatorname{dist}(F, G)$. Since $r < 1$, this forces $\operatorname{dist}(F, G) = 0$, so that $F = G$, and the proof of Theorem 2.9 is complete.

Under an additional technical condition, one can calculate the precise Hausdorff dimension of the self-similar set F. Loosely speaking, the restriction holds if the sets $S_1(F), \ldots, S_m(F)$ do not overlap too much. Indeed, if these sets were disjoint, then we could argue that

$$m_\alpha(F) = \sum_{j=1}^{m} m_\alpha(S_j(F)).$$

Since each S_j scales by r, we would then have $m_\alpha(S_j(F)) = r^\alpha m_\alpha(F)$. Hence

$$m_\alpha(F) = mr^\alpha m_\alpha(F).$$

If $m_\alpha(F)$ were finite, then we would have that $mr^\alpha = 1$; thus

$$\alpha = \frac{\log m}{\log 1/r}.$$

The restriction we impose is as follows. We say that the similarities $S_1, \ldots, S_m$ are **separated** if there is an bounded open set $\mathcal{O}$ so that

$$\mathcal{O} \supset S_1(\mathcal{O}) \cup \cdots \cup S_m(\mathcal{O}),$$

and the $S_j(\mathcal{O})$ are disjoint. It is not assumed that $\mathcal{O}$ contains F.

Theorem 2.12 *Suppose $S_1, S_2, \ldots, S_m$ are m separated similarities with the common ratio r that satisfies $0 < r < 1$. Then the set F has Hausdorff dimension equal to $\log m / \log(1/r)$.*

Observe first that when F is the Cantor set we may take $\mathcal{O}$ to be the open unit interval, and note that we have already proved that its dimension is $\log 2 / \log 3$. For the Sierpinski triangle the open unit triangle will do, and $\dim \mathcal{S} = \log 3 / \log 2$. In the example of the Cantor dust the open unit square works, and $\dim \mathcal{D} = \log m / \log \mu^{-1}$. Finally, for the von Koch curve we may take the interior of the triangle pictured in Figure 7, and we will have $\dim \mathcal{K} = \log 4 / \log 3$.

We now turn to the proof of Theorem 2.12, which will follow the same approach used in the case of the Sierpinski triangle. If $\alpha = \log m / \log(1/r)$, we claim that $m_\alpha(F) < \infty$, hence $\dim F \leq \alpha$. Moreover, this inequality holds even without the separation assumption. Indeed, recall that

$$F_k = \tilde{S}^k(B),$$

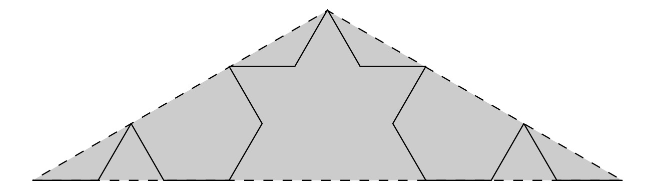


Figure 7. Open set in the separation of the von Koch similarities

and $\tilde{S}^k(B)$ is the union of m^k sets of diameter less than cr^k (with $c =$ diam B), each of the form

$$S_{n_1} \circ S_{n_2} \circ \cdots \circ S_{n_k}(B), \qquad \text{where } 1 \leq n_i \leq m \text{ and } 1 \leq i \leq k.$$

Consequently, if $cr^k \leq \delta$, then

$$\begin{aligned}\mathcal{H}_\alpha^\delta(F) &\leq \sum_{n_1,\ldots,n_k} (\text{diam } S_{n_1} \circ \cdots \circ S_{n_k}(B))^\alpha \\ &\leq c' m^k r^{\alpha k} \\ &\leq c',\end{aligned}$$

since $mr^\alpha = 1$, because $\alpha = \log m / \log(1/r)$. Since c' is independent of δ, we get $m_\alpha(F) \leq c'$.

To prove $m_\alpha(F) > 0$, we now use the separation condition. We argue in parallel with the earlier calculation of the Hausdorff dimension of the Sierpinski triangle.

Fix a point $\overline{x}$ in F. We define *the* "vertices" of the kth generation as the m^k points that lie in F and are given by

$$S_{n_1} \circ \cdots \circ S_{n_k}(\overline{x}), \qquad \text{where } 1 \leq n_1 \leq m, \ldots, 1 \leq n_k \leq m.$$

Each vertex is labeled by $(n_1, \ldots, n_k)$. Vertices need not be distinct, so they are counted with their multiplicities.

Similarly, we define *the* "open sets" of the kth generation to be the m^k sets given by

$$S_{n_1} \circ \cdots \circ S_{n_k}(\mathcal{O}), \qquad \text{where } 1 \leq n_1 \leq m, \ldots, 1 \leq n_k \leq m,$$

and where $\mathcal{O}$ is fixed and chosen to satisfy the separation condition. Such open sets are again labeled by multi-indices $(n_1, n_2, \ldots, n_k)$ with $1 \leq n_j \leq m$, $1 \leq j \leq k$.

Then the open sets of the kth generation are disjoint, since those of the first generation are disjoint. Moreover if $k \geq \ell$, each open set of the ℓth generation contains $m^{k-\ell}$ open sets of the kth generation.

Suppose v is a vertex of the kth generation, and let $\mathcal{O}(v)$ denote the open set in the kth generation which is associated to v, that is, v and $\mathcal{O}(v)$ carry the same label $(n_1, n_2, \ldots, n_k)$. Since $\overline{x}$ is at a fixed distance from the original open set $\mathcal{O}$, and $\mathcal{O}$ has a finite diameter, we find that

(a) $d(v, \mathcal{O}(v)) \leq cr^k$.

(b) $c'r^k \leq \text{diam } \mathcal{O}(v) \leq cr^k$.

As in the case of the Sierpinski triangle, it suffices to prove that if $\mathcal{B} = \{B_j\}_{j=1}^N$ is a finite collection of balls whose diameters are less than δ and whose union covers F, then

$$\sum_{j=1}^{N} (\text{diam } B_j)^\alpha \geq c > 0.$$

Suppose we have such a covering by balls, and choose k so that

$$r^k \leq \min_{1 \leq j \leq N} \text{diam } B_j < r^{k-1}.$$

Lemma 2.13 *Suppose B is a ball in the covering $\mathcal{B}$ that satisfies*

$$r^\ell \leq \text{diam } B < r^{\ell-1} \quad \textit{for some } \ell \leq k.$$

Then B contains at most $cm^{k-\ell}$ vertices of the kth generation.

Proof. If v is a vertex of the kth generation with $v \in B$, and $\mathcal{O}(v)$ denotes the corresponding open set of the kth generation, then, for some fixed dilate B^* of B, properties (a) and (b) above guarantee that $\mathcal{O}(v) \subset B^*$, and B^* also contains the open set of generation ℓ that contains $\mathcal{O}(v)$.

Since B^* has volume $cr^{d\ell}$, and each open set in the ℓth generation has volume $\approx r^{d\ell}$ (by property (b) above), B^* can contain at most c open sets of generation ℓ. Hence B^* contains at most $cm^{k-\ell}$ open sets of the kth generation. Consequently, B can contain at most $cm^{k-\ell}$ vertices of the kth generation, and the lemma is proved.

For the final argument, let N_ℓ denote the number of balls in $\mathcal{B}$ so that

$$r^\ell \leq \text{diam } B_j \leq r^{\ell-1}.$$

By the lemma, we see that the total number of vertices of the kth generation that can be covered by the collection $\mathcal{B}$ can be no more than

$c\sum_\ell N_\ell m^{k-\ell}$. Since all m^k vertices of the kth generation belong to F, we must have $c\sum_\ell N_\ell m^{k-\ell} \geq m^k$, and hence

$$\sum_\ell N_\ell m^{-\ell} \geq c.$$

The definition of α gives $r^{\ell\alpha} = m^{-\ell}$, and therefore

$$\sum_{j=1}^{N} (\text{diam } B_j)^\alpha \geq \sum_\ell N_\ell r^{\ell\alpha} \geq c,$$

and the proof of Theorem 2.12 is complete.

3 Space-filling curves

The year 1890 heralded an important discovery: Peano constructed a continuous curve that filled an entire square in the plane. Since then, many variants of his construction have been given. We shall describe here a construction that has the feature of elucidating an additional significant fact. It is that from the point of measure theory, speaking broadly, the unit interval and unit square are "isomorphic."

Theorem 3.1 *There exists a curve $t \mapsto \mathcal{P}(t)$ from the unit interval to the unit square with the following properties:*

(i) *$\mathcal{P}$ maps $[0,1]$ to $[0,1] \times [0,1]$ continuously and surjectively.*

(ii) *$\mathcal{P}$ satisfies a Lipschitz condition of exponent $1/2$, that is,*

$$|\mathcal{P}(t) - \mathcal{P}(s)| \leq M|t-s|^{1/2}.$$

(iii) *The image under $\mathcal{P}$ of any sub-interval $[a,b]$ is a compact subset of the square of (two-dimensional) Lebesgue measure exactly $b-a$.*

The third conclusion can be elaborated further.

Corollary 3.2 *There are subsets $Z_1 \subset [0,1]$ and $Z_2 \subset [0,1] \times [0,1]$, each of measure zero, such that $\mathcal{P}$ is bijective from*

$$[0,1] - Z_1 \quad \textit{to} \quad [0,1] \times [0,1] - Z_2$$

and measure preserving. In other words, E is measurable if and only if $\mathcal{P}(E)$ is measurable, and

$$m_1(E) = m_2(\mathcal{P}(E)).$$

Here m_1 and m_2 denote the Lebesgue measures in $\mathbb{R}^1$ and $\mathbb{R}^2$, respectively.

We shall call the function $t \mapsto \mathcal{P}(t)$ the **Peano mapping**. Its image is called the **Peano curve**.

Several observations help clarify the nature of the conclusions of the theorem. Suppose that $F : [0,1] \to [0,1] \times [0,1]$ is continuous and surjective. Then:

(a) F cannot be Lipschitz of exponent $\gamma > 1/2$. This follows at once from Lemma 2.2, which states that

$$\dim\ F([0,1]) \leq \frac{1}{\gamma} \dim[0,1],$$

so that $2 \leq 1/\gamma$ as desired.

(b) F cannot be injective. Indeed, if this were the case, then the inverse G of F would exist and would be continuous. Given any two points $a \neq b$ in $[0,1]$, we would get a contradiction by looking at two distinct curves in the square that join $F(a)$ and $F(b)$, since the image of these two curves under G would have to intersect at points between a and b. In fact, given any open disc D in the square, there always exists $x \in D$ so that $F(t) = F(s) = x$ yet $t \neq s$.

The proof of Theorem 3.1 will follow from a careful study of a natural class of mappings that associate sub-squares in $[0,1] \times [0,1]$ to sub-intervals in $[0,1]$. This implements the approach implicit in Hilbert's iterative procedure, which he set forth in the first three stages in Figure 8.

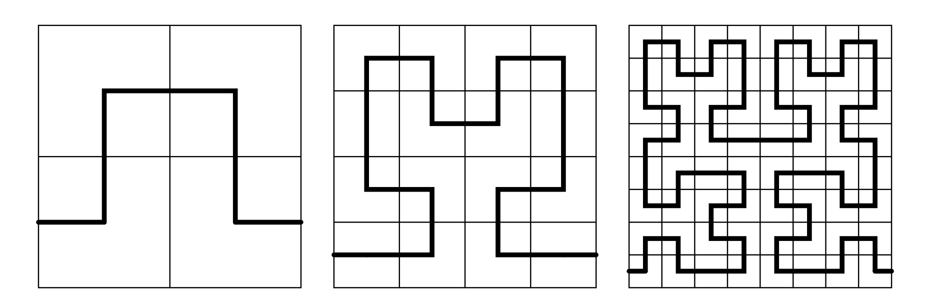


Figure 8. Construction of the Peano curve

We turn now to the study of the general class of mappings.

3.1 Quartic intervals and dyadic squares

The **quartic intervals** arise when $[0, 1]$ is successively sub-divided by powers of 4. For instance, the first generation quartic intervals are the closed intervals

$$I_1 = [0, 1/4], \quad I_2 = [1/4, 1/2], \quad I_3 = [1/2, 3/4], \quad I_4 = [3/4, 1].$$

The second generation quartic intervals are obtained by sub-dividing each interval of the first generation by 4. Hence there are $16 = 4^2$ quartic intervals of the second generation. In general, there are 4^k quartic intervals of the kth generation, each of the form $[\frac{\ell}{4^k}, \frac{\ell+1}{4^k}]$, where ℓ is integral with $0 \leq \ell < 4^k$.

A **chain** of quartic intervals is a decreasing sequence of intervals

$$I^1 \supset I^2 \supset \cdots \supset I^k \supset \cdots ,$$

where I^k is a quartic interval of the kth generation (hence $|I^k| = 4^{-k}$).

Proposition 3.3 *Chains of quartic intervals satisfy the following properties:*

(i) *If $\{I^k\}$ is a chain of quartic intervals, then there exists a unique $t \in [0, 1]$ such that $t \in \bigcap_k I^k$.*

(ii) *Conversely, given $t \in [0, 1]$, there is a chain $\{I^k\}$ of quartic intervals such that $t \in \bigcap_k I^k$.*

(iii) *The set of t for which the chain in part* (ii) *is not unique is a set of measure zero (in fact, this set is countable).*

Proof. Part (i) follows from the fact that $\{I^k\}$ is a decreasing sequence of compact sets whose diameters go to 0.

For part (ii), we fix t and note that for each k there exists at least one quartic interval I^k with $t \in I^k$. If t is of the form $\ell/4^k$, where $0 < \ell < 4^k$, then there are exactly two quartic intervals of the kth generation that contain t. Hence, the set of points for which the chain is not unique is precisely the set of **dyadic rationals**

$$\frac{\ell}{4^k}, \qquad \text{where } 1 \leq k, \text{ and } 0 < \ell < 4^k.$$

Note that of course, these fractions are the same as those of the form $\ell'/2^{k'}$ with $0 < \ell' < 2^{k'}$. This set is countable, hence has measure 0.

It is clear that each chain $\{I^k\}$ of quartic intervals can be represented naturally by a string $.a_1a_2\cdots a_k\cdots$, where each a_k is either $0, 1, 2$, or 3. Then the point t corresponding to this chain is given by

$$t = \sum_{k=1}^{\infty} \frac{a_k}{4^k}.$$

The points where ambiguity occurs are precisely those where $a_k = 3$ for all sufficiently large k, or equivalently where $a_k = 0$ for all sufficiently large k.

Part of our description of the Peano mapping will follow from associating to each quartic interval a dyadic square. These **dyadic squares** are obtained by sub-dividing the unit square $[0,1] \times [0,1]$ in the plane by successively bisecting the sides.

For instance, dyadic squares of the first generation arise from bisecting the sides of the unit square. This yields four closed squares S_1, S_2, S_3 and S_4, each of side length $1/2$ and area $|S_i| = 1/4$, for $i = 1, \ldots, 4$.

The dyadic squares of the second generation are obtained by bisecting each dyadic square of the first generation, and so on. In general, there are 4^k squares of the kth generation, each of side length $1/2^k$ and area $1/4^k$.

A **chain** of dyadic squares is a decreasing sequence of squares

$$S^1 \supset S^2 \supset \cdots \supset S^k \supset \cdots,$$

where S^k is a dyadic square of the kth generation.

Proposition 3.4 *Chains of dyadic squares have the following properties:*

(i) *If $\{S^k\}$ is a chain of dyadic squares, then there exists a unique $x \in [0,1] \times [0,1]$ such that $x \in \bigcap_k S^k$.*

(ii) *Conversely, given $x \in [0,1] \times [0,1]$, there is a chain $\{S^k\}$ of dyadic squares such that $x \in \bigcap_k S^k$.*

(iii) *The set of x for which the chain in part* (ii) *is not unique is a set of measure zero.*

In this case, the set of ambiguities consists of all points (x_1, x_2) where one of the coordinates is a dyadic rational. Geometrically, this set is the (countable) union of vertical and horizontal segments in $[0,1] \times [0,1]$ determined by the grid of dyadic rationals. This set has measure zero.

Moreover, each chain of dyadic squares can be represented by a string $.b_1 b_2 \cdots$, where each b_k is either $0, 1, 2$ or 3. Then

$$x = \sum_{k=1}^{\infty} \frac{\overline{b}_k}{2^k}, \tag{1}$$

where

$$\begin{aligned} \overline{b}_k &= (0,0) \quad \text{if } b_k = 0, \\ \overline{b}_k &= (0,1) \quad \text{if } b_k = 1, \\ \overline{b}_k &= (1,0) \quad \text{if } b_k = 2, \\ \overline{b}_k &= (1,1) \quad \text{if } b_k = 3. \end{aligned}$$

3.2 Dyadic correspondence

A **dyadic correspondence** is a mapping Φ from quartic intervals to dyadic squares that satisfies:

(1) Φ is bijective.

(2) Φ respects generations.

(3) Φ respects inclusion.

By (2), we mean that if I is a quartic interval of the kth generation, then $\Phi(I)$ is a dyadic square of the kth generation. By (3), we mean that if $I \subset J$, then $\Phi(I) \subset \Phi(J)$.

For example, the trivial, or standard correspondence assigns to the string $.a_1 a_2 \cdots$ the string $.b_1 b_2 \cdots$ with $b_k = a_k$.

Given a dyadic correspondence Φ, the **induced mapping** Φ^* maps $[0,1]$ to $[0,1] \times [0,1]$ and is given as follows. If $\{t\} = \bigcap I^k$ where $\{I^k\}$ is a chain of quartic intervals, then, since $\{\Phi(I^k)\}$ is a chain of dyadic squares, we may let

$$\Phi^*(t) = x = \bigcap \Phi(I^k).$$

We note that Φ^* is well-defined except on a (countable) set of measure zero, (those points t that are represented by more than one quartic chain.)

A moment's reflection will show that if I' is a quartic interval of the kth generation, then the images $\Phi^*(I') = \{\Phi^*(t),\ t \in I'\}$, comprise the dyadic square of the kth generation $\Phi(I')$. Thus $\Phi^*(I') = \Phi(I')$, and hence $m_1(I') = m_2(\Phi^*(I'))$.

Theorem 3.5 *Given a dyadic correspondence Φ, there exist sets $Z_1 \subset [0,1]$ and $Z_2 \subset [0,1] \times [0,1]$, each of measure zero, so that:*

(i) *Φ^* is a bijection on $[0,1]-Z_1$ to $[0,1]\times[0,1]-Z_2$.*

(ii) *E is measurable if and only if $\Phi^*(E)$ is measurable.*

(iii) *$m_1(E)=m_2(\Phi^*(E))$.*

Proof. First, let $\mathcal{N}_1$ denote the collection of chains of those quartic intervals arising in (iii) of Proposition 3.3, those for which the points in $I=[0,1]$ are not uniquely representable. Similarly, let $\mathcal{N}_2$ denote the collection of chains of those dyadic squares for which the corresponding points in the square $I\times I$ are not uniquely representable.

Since Φ is a bijection from chains of quartic intervals to chains of dyadic squares, it is also a bijection from $\mathcal{N}_1\cup\Phi^{-1}(\mathcal{N}_2)$ to $\Phi(\mathcal{N}_1)\cup\mathcal{N}_2$, and hence also of their complements. Let Z_1 be the subset of I consisting of all points in I that can be represented (according to (i) of Proposition 3.3) by the chains in $\mathcal{N}_1\cup\Phi^{-1}(\mathcal{N}_2)$, and let Z_2 be the set of points in the square that can be represented by dyadic squares in $\Phi(\mathcal{N}_1)\cup\mathcal{N}_2$. Then Φ^*, the induced mapping, is well-defined on $I-Z_1$, and gives a bijection of $I-Z_1$ to $(I\times I)-Z_2$. To prove that both Z_1 and Z_2 have measure zero, we invoke the following lemma. We suppose $\{f_k\}_{k=1}^{\infty}$ is a fixed given sequence, with each f_k either $0,1,2$, or 3.

Lemma 3.6 *Let*

$$E_0=\{x=\sum_{k=1}^{\infty}a_k/4^k,\quad \text{where } a_k\neq f_k \text{ for all sufficiently large } k\}.$$

Then $m(E_0)=0$.

Indeed, if we fix r, then $m(\{x:\ a_r\neq f_r\})=3/4$, and

$$m(\{x:\ a_r\neq f_r \text{ and } a_{r+1}\neq f_{r+1}\})=(3/4)^2,\qquad \text{etc.}$$

Thus $m(\{x:\ a_k\neq f_k,\ \text{all } k\geq r\})=0$, and E_0 is a countable union of such sets, from which the lemma follows.

There is a similar statement for points in the square $S=I\times I$ in terms of the representation (1).

Note that as a result the set of points in I corresponding to chains in $\mathcal{N}_1$ form a set of measure zero. In fact, we may use the lemma for the sequence for which $f_k=1$, for all k, since the elements of $\mathcal{N}_1$ correspond to sequences $\{a_k\}$ with $a_k=0$ for all sufficiently large k, or $a_k=3$ for all sufficiently large k.

Similarly, the points in the square S corresponding to $\mathcal{N}_2$ form a set of measure zero. To see this, take for example $f_k=1$ for k odd, and $f_k=2$

for k even, and note that $\mathcal{N}_2$ corresponds to all sequences $\{a_k\}$ where one of the following four exclusive alternatives holds for all sufficiently large k: either a_k is 0 or 1; or a_k is 2 or 3; or a_k is 0 or 2; or a_k is 1 or 3. By similar reasoning the points $\Phi^{-1}(\mathcal{N}_2)$ and $\Phi(\mathcal{N}_1)$ form sets of measure zero in I and $I \times I$ respectively.

We now turn to the proof that Φ^* (which is a bijection from $I - Z_1$ to $(I \times I) - Z_2$) is measure preserving. For this it is useful to recall Theorem 1.4 in Chapter 1, whereby any open set $\mathcal{O}$ in the unit interval I can be realized as a countable union $\bigcup_{j=1}^{\infty} I_j$, where each I_j is a closed interval and the I_j have disjoint interiors. Moreover, an examination of the proof shows that the intervals can be taken to be dyadic, that is, of the form $[\ell/2^j, (\ell+1)/2^j]$, for appropriate integers ℓ and j. Further, such an interval is itself a quartic interval if j is even, $j = 2k$, or the union of two quartic intervals $[(2\ell)/2^{2k}, (2\ell+1)/2^{2k}]$ and $[(2\ell+1)/2^{2k}, (2\ell+2)/2^{2k}]$, if j is odd, $j = 2k-1$. Thus any open set in I can be given as a union of quartic intervals whose interiors are disjoint. Similarly, any open set in the square $I \times I$ is a union of dyadic squares whose interiors are disjoint.

Now let E be any set of measure zero in $I - Z_1$ and $\epsilon > 0$. Then we can cover $E \subset \bigcup_j I_j$, where I_j are quartic intervals and $\sum_j m_1(I_j) < \epsilon$. Because $\Phi^*(E) \subset \bigcup_j \Phi^*(I_j)$, then

$$m_2(\Phi^*(E)) \leq \sum m_2(\Phi^*(I_j)) = \sum m_1(I_j) < \epsilon.$$

Thus $\Phi^*(E)$ is measurable and $m_2(\Phi^*(E)) = 0$. Similarly, $(\Phi^*)^{-1}$ maps sets of measure zero in $(I \times I) - Z_2$ to sets of measure zero in I.

Now the argument above also shows that if $\mathcal{O}$ is any open set in I, then $\Phi^*(\mathcal{O} - Z_1)$ is measurable, and $m_2(\Phi^*(\mathcal{O} - Z_1)) = m_1(\mathcal{O})$. Thus this identity goes over to G_δ sets in I. Since any measurable set differs from a G_δ set by a set of measure zero, we see that we have established that $m_2(\Phi^*(E)) = m_1(E)$ for any measurable subset of E of $I - Z_1$. The same argument can be applied to $(\Phi^*)^{-1}$, and this completes the proof of the theorem.

The Peano mapping will be obtained as Φ^* for a special correspondence Φ.

3.3 Construction of the Peano mapping

The particular dyadic correspondence we now present provides us with the steps to follow when tracing the approximations of the Peano curve. The main idea behind its construction is that as we go from one quartic interval in the kth generation to the next quartic interval in the same

generation, we move from a dyadic square of the kth generation to another square of the kth generation that shares a common side.

More precisely, we say that two quartic intervals in the same generation are **adjacent** if they share a point in common. Also, two squares in the same generation are **adjacent** if they share a side in common.

Lemma 3.7 *There is a unique dyadic correspondence Φ so that:*

(i) *If I and J are two adjacent intervals of the same generation, then $\Phi(I)$ and $\Phi(J)$ are two adjacent squares (of the same generation).*

(ii) *In generation k, if I_- is the left-most interval and I_+ the right-most interval, then $\Phi(I_-)$ is the left-lower square and $\Phi(I_+)$ is the right-lower square.*

Part (ii) of the lemma is illustrated in Figure 9.

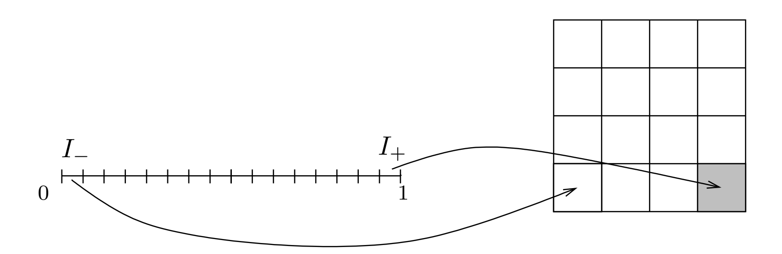


Figure 9. Special dyadic correspondence

Given a square S and its four immediate sub-squares, an acceptable **traverse** is an ordering of the sub-squares S_1, S_2, S_3, and S_4, so that S_j and S_{j+1} are adjacent for $j = 1, 2, 3$. With such an ordering, we note that if we color S_1 white, and then alternate black and white, the square S_3 is also white, while S_2 and S_4 are black. The important point to remember is that if the first square in a traverse is white, then the last square is black.

The key observation is the following. Suppose we are given a square S, and a side σ of S. If S_1 is any of the immediate four sub-squares in S, then there exists a unique traverse S_1, S_2, S_3, and S_4 so that the last square S_4 has a side in common with σ. With the initial square S_1 in the lower-left corner of S, the four possibilities which correspond to the four choices of σ, are illustrated in Figure 10.

We may now begin the inductive description of the dyadic correspondence satisfying the conditions in the lemma. On quartic intervals of the first generation we assign the square $S_j = \Phi(I_j)$, as pictured in Figure 11.

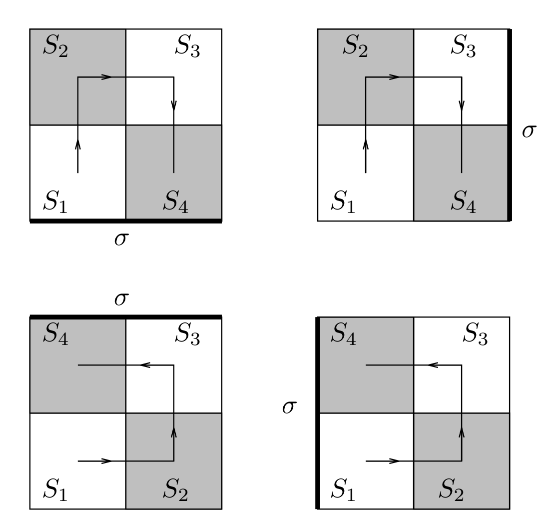


Figure 10. Traverses

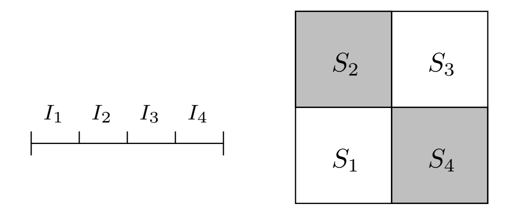


Figure 11. Initial step of the correspondence

Now suppose Φ has been defined for all quartic intervals of generation less than or equal to k. We now write the intervals in generation k in increasing order as $I_1, \ldots, I_{4^k}$, and let $S_j = \Phi(I_j)$. We then divide I_1 into four quartic intervals of generation $k+1$ and denote them by $I_{1,1}$, $I_{1,2}$, $I_{1,3}$, and $I_{1,4}$, where the intervals are chosen in increasing order.

Then, we assign to each interval $I_{1,j}$ a dyadic square $\Phi(I_{1,j}) = S_j$ of generation $k+1$ contained in S_1 so that:

(a) $S_{1,1}$ is the lower-left sub-square of S_1,

(b) $S_{1,4}$ touches the side that S_1 shares with S_2,

(c) $S_{1,1}$, $S_{1,2}$, $S_{1,3}$, and $S_{1,4}$ is a traverse.

This is possible, since the induction hypothesis guarantees that S_2 is adjacent to S_1.

This settles the assignments for the sub-squares of S_1, so we now turn our attention to S_2. Let $I_{2,1}$, $I_{2,2}$, $I_{2,3}$, and $I_{2,4}$ denote the quartic intervals of generation $k+1$ in I_2, written in increasing order. First, we take $S_{2,1} = \Phi(I_{2,1})$ to be the sub-square of S_2 which is adjacent to $S_{1,4}$. This can be done because $S_{1,4}$ touches S_2 by construction. Note that we leave S_1 from a black square ($S_{1,4}$), and enter S_2 in a white square ($S_{2,1}$). Since S_3 is adjacent to S_2, we may now find a traverse $S_{2,1}$, $S_{2,2}$, $S_{2,3}$ and $S_{2,4}$ so that $S_{2,4}$ touches S_3.

We may then repeat this process in each interval I_j and square S_j, $j = 3, \dots, 4^k$. Note that at each stage the square $S_{j,1}$ (the "entering" square) is white, while $S_{j,4}$ (the "exiting" square) is black.

In the final step, the induction hypothesis guarantees that S_{4^k} is the lower-right corner square. Moreover, since S_{4^k-1} must be adjacent to S_{4^k} it must be either above it, or to the left of it, so we enter a square of the $(k+1)^{\text{st}}$ generation along an upper or left side. The entering square is a white square, and we traverse to the lower right corner sub-square of S_{4^k}, which is a black square.

This concludes the inductive step, hence the proof of Lemma 3.7.

We may now begin the actual description of the Peano curve. For each generation k we construct a polygonal line which consists of vertical and horizontal line segments connecting the centers of consecutive squares. More precisely, let Φ denote the dyadic correspondence in Lemma 3.7, and let $S_1, \dots, S_{4^k}$ be the squares of the kth generation ordered according to Φ, that is, $\Phi(I_j) = S_j$. Let t_j denote the middle point of I_j,

$$t_j = \frac{j - \frac{1}{2}}{4^k} \qquad \text{for } j = 1, \dots, 4^k.$$

Let x_j be the center of the square S_j, and define

$$\mathcal{P}_k(t_j) = x_j.$$

Also set

$$\mathcal{P}_k(0) = (0, 1/2^{k+1}) = x_0 \qquad \text{where } t_0 = 0,$$

and

$$\mathcal{P}_k(1) = (1, 1/2^{k+1}) = x_{4^k+1} \qquad \text{where } t_{4^k+1} = 1.$$

Then, we extend $\mathcal{P}_k(t)$ to the unit interval $0 \le t \le 1$ by linearity along the sub-intervals determined by the division points $t_0, \dots, t_{4^k+1}$.

Note that the distance $|x_j - x_{j+1}| = 1/2^k$, while $|t_j - t_{j+1}| = 1/4^k$ for $0 \leq j \leq 4^k$. Also

$$|x_1 - x_0| = |x_{4^k} - x_{4^k+1}| = \frac{1}{2 \cdot 2^k},$$

while

$$|t_1 - t_0| = |t_{4^k} - t_{4^k+1}| = \frac{1}{2 \cdot 4^k}.$$

Therefore $\mathcal{P}_k'(t) = 4^k 2^{-k} = 2^k$ except when $t = t_j$.

As a result,

$$|\mathcal{P}_k(t) - \mathcal{P}_k(s)| \leq 2^k |t - s|.$$

However,

$$|\mathcal{P}_{k+1}(t) - \mathcal{P}_k(t)| \leq \sqrt{2}\, 2^{-k},$$

because when $\ell/4^k \leq t \leq (\ell+1)/4^k$, then $\mathcal{P}_{k+1}(t)$ and $\mathcal{P}_k(t)$ belong to the same dyadic square of generation k.

Therefore the limit

$$\mathcal{P}(t) = \lim_{k\to\infty} \mathcal{P}_k(t) = \mathcal{P}_1(t) + \sum_{j=1}^{\infty} \mathcal{P}_{j+1}(t) - \mathcal{P}_j(t)$$

exists, and defines a continuous function in view of the uniform convergence. By Lemma 2.8 we conclude that

$$|\mathcal{P}(t) - \mathcal{P}(s)| \leq M|t - s|^{1/2},$$

and $\mathcal{P}$ satisfies a Lipschitz condition of exponent of 1/2.

Moreover, each $\mathcal{P}_k(t)$ visits each dyadic square of generation k as t ranges in $[0, 1]$. Hence $\mathcal{P}$ is dense in the unit square, and by continuity we find that $t \mapsto \mathcal{P}(t)$ is a surjection.

Finally, to prove the measure preserving property of $\mathcal{P}$, it suffices to establish $\mathcal{P} = \Phi^*$.

Lemma 3.8 *If Φ is the dyadic correspondence in Lemma 3.7, then $\Phi^*(t) = \mathcal{P}(t)$ for every $0 \leq t \leq 1$.*

Proof. First, we observe that $\Phi^*(t)$ is unambiguously defined for every t. Indeed, suppose $t \in \bigcap_k I^k$ and $t \in \bigcap_k J^k$ are two chains of quartic intervals; then I^k and J^k must be adjacent for sufficiently large

k. Thus $\Phi(I^k)$ and $\Phi(J^k)$ must be adjacent squares for all sufficiently large k. Hence

$$\bigcap_k \Phi(I^k) = \bigcap_k \Phi(J^k).$$

Next, directly from our construction we have

$$\bigcap_k \Phi(I^k) = \lim \mathcal{P}_k(t) = \mathcal{P}(t).$$

This gives the desired conclusion.

The argument also shows that $\mathcal{P}(I) = \Phi(I)$ for any quartic interval I. Now recall that any interval (a, b) can be written as $\bigcup_j I_j$, where the I_j are quartic intervals with disjoint interiors. Because $\mathcal{P}(I_j) = \Phi(I_j)$, these are then dyadic squares with disjoint interiors. Since $\mathcal{P}(a, b) = \bigcup_k \mathcal{P}(I_j)$, we have

$$m_2(\mathcal{P}(a,b)) = \sum_{j=1}^{\infty} m_2(\mathcal{P}(I_j)) = \sum_{j=1}^{\infty} m_2(\Phi(I_j)) = \sum_{j=1}^{\infty} m_1(I_j) = m_1(a,b).$$

This proves conclusion (iii) of Theorem 3.1. The other conclusions having already been established, we need only note that the corollary is contained in Theorem 3.5.

As a result, we conclude that $t \mapsto \mathcal{P}(t)$ also induces a measure preserving mapping from $[0, 1]$ to $[0, 1] \times [0, 1]$. This concludes the proof of Theorem 3.1.

4* Besicovitch sets and regularity

We begin by presenting a surprising regularity property enjoyed by all measurable subsets (of finite measure) of $\mathbb{R}^d$ when $d \geq 3$. As we shall see, the fact that the corresponding phenomenon does not hold for $d = 2$ is due to the existence of a remarkable set that was discovered by Besicovitch. A construction of a set of this kind will be detailed in Section 4.4.

We first fix some notation. For each unit vector γ on the sphere, $\gamma \in S^{d-1}$, and each $t \in \mathbb{R}$ we consider the **plane** $\mathcal{P}_{t,\gamma}$, which is defined as the $(d-1)$-dimensional affine hyperplane perpendicular to γ and of "signed distance" t from the origin.[1] The plane $\mathcal{P}_{t,\gamma}$ is given by

$$\mathcal{P}_{t,\gamma} = \{x \in \mathbb{R}^d : \ x \cdot \gamma = t\}.$$

[1] Note that there are two planes perpendicular to γ and of distance $|t|$ from the origin; this accounts for the fact that t may be either positive or negative.

We observe that each $\mathcal{P}_{t,\gamma}$ carries a natural $(d-1)$ Lebesgue measure, denoted by m_{d-1}. In fact, if we complete γ to an orthonormal basis $e_1, e_2, \ldots, e_{d-1}, \gamma$ of $\mathbb{R}^d$, then we can write any $x \in \mathbb{R}^d$ in terms of the corresponding coordinates as $x = x_1e_1 + x_2e_2 + \cdots + x_d\gamma$. When we set $x \in \mathbb{R}^d = \mathbb{R}^{d-1} \times \mathbb{R}$ with $(x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1}$, $x_d \in \mathbb{R}$, then the measure m_{d-1} on $\mathcal{P}_{t,\gamma}$ is the Lebesgue measure on $\mathbb{R}^{d-1}$. This definition of m_{d-1} is independent of the choice of orthonormal vectors $e_1, e_2, \ldots, e_{d-1}$, since Lebesgue measure is invariant under rotations. (See Problem 4, Chapter 2, or Exercise 26, Chapter 3.)

With these preliminaries out of the way, we define for each subset $E \subset \mathbb{R}^d$ the **slice** of E cut out by the plane $\mathcal{P}_{t,\gamma}$ as

$$E_{t,\gamma} = E \cap \mathcal{P}_{t,\gamma}.$$

We now consider the slices $E_{t,\gamma}$ as t varies, where E is measurable and γ is fixed. (See Figure 12.)

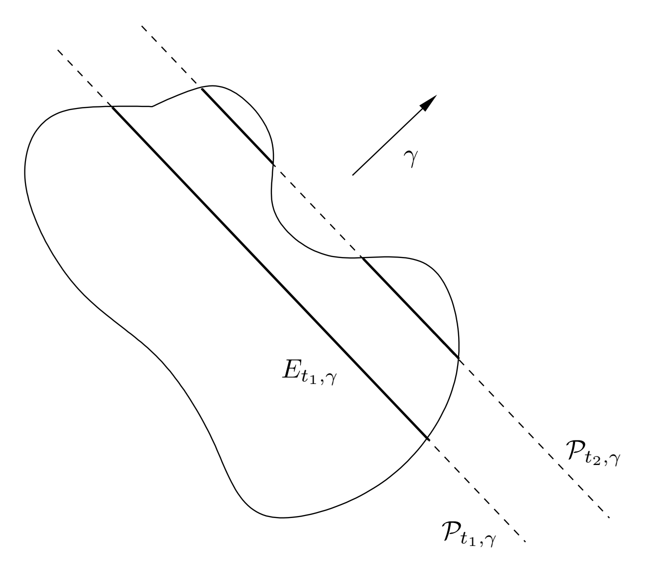


Figure 12. The slices $E \cap \mathcal{P}_{t,\gamma}$ as t varies

We observe that for almost every t the set $E_{t,\gamma}$ is m_{d-1} measurable and, moreover, $m_{d-1}(E_{t,\gamma})$ is a measurable function of t. This is a direct consequence of Fubini's theorem and the above decomposition, $\mathbb{R}^d = \mathbb{R}^{d-1} \times \mathbb{R}$. In fact, so long as the direction γ is pre-assigned, not much more can be said in general about the function $t \mapsto m_{d-1}(E_{t,\gamma})$.

However, when $d \geq 3$ the nature of the function is dramatically different for "most" γ. This is contained in the following theorem.

Theorem 4.1 *Suppose E is of finite measure in $\mathbb{R}^d$, with $d \geq 3$. Then for almost every $\gamma \in S^{d-1}$:*

(i) *$E_{t,\gamma}$ is measurable for all $t \in \mathbb{R}$.*

(ii) *$m_{d-1}(E_{t,\gamma})$ is continuous in $t \in \mathbb{R}$.*

Moreover, the function of t defined by $\mu(t,\gamma) = m_{d-1}(E_{t,\gamma})$ satisfies a Lipschitz condition with exponent α for any α with $0 < \alpha < 1/2$.

The almost everywhere assertion is with respect to the natural measure $d\sigma$ on S^{d-1} that arises in the polar coordinate formula in Section 3.2 of the previous chapter.

We recall that a function f is Lipschitz with exponent α if

$$|f(t_1) - f(t_2)| \leq A|t_1 - t_2|^{\alpha} \quad \text{for some } A.$$

A significant part of (i) is that for a.e. γ, the slice $E_{t,\gamma}$ is measurable for *all* values of the parameter t. In particular, one has the following.

Corollary 4.2 *Suppose E is a set of measure zero in $\mathbb{R}^d$ with $d \geq 3$. Then, for almost every $\gamma \in S^{d-1}$, the slice $E_{t,\gamma}$ has zero measure for all $t \in \mathbb{R}$.*

The fact that there is no analogue of this when $d = 2$ is a consequence of the existence of a **Besicovitch set**, (also called a "Kakeya set"), which is defined as a set that satisfies the three conditions in the theorem below.

Theorem 4.3 *There exists a set $\mathcal{B}$ in $\mathbb{R}^2$ that:*

(i) *is compact,*

(ii) *has Lebesgue measure zero,*

(iii) *contains a translate of every unit line segment.*

Note that with $F = \mathcal{B}$ and $\gamma \in S^1$ one has $m_1(F \cap \mathcal{P}_{t_0,\gamma}) \geq 1$ for some t_0. If $m_1(F \cap \mathcal{P}_{t,\gamma})$ were continuous in t, then this measure would be strictly positive for an interval in t containing t_0, and thus we would have $m_2(F) > 0$, by Fubini's theorem. This contradiction shows that the analogue of Theorem 4.1 cannot hold for $d = 2$.

While the set $\mathcal{B}$ has zero two-dimensional measure, this assertion cannot be improved by replacing this measure by α-dimensional Hausdorff measure, with $\alpha < 2$.

Theorem 4.4 *Suppose F is any set that satisfies the conclusions* (i) *and* (iii) *of Theorem* 4.3. *Then F has Hausdorff dimension* 2.

4.1 The Radon transform

Theorems 4.1 and 4.4 will be derived by an analysis of the regularity properties of the Radon transform $\mathcal{R}$. The operator $\mathcal{R}$ arises in a number of problems in analysis, and was already considered in Chapter 6 of Book I.

For an appropriate function f on $\mathbb{R}^d$, the **Radon transform** of f is defined by

$$\mathcal{R}(f)(t,\gamma) = \int_{\mathcal{P}_{t,\gamma}} f.$$

The integration is performed over the plane $\mathcal{P}_{t,\gamma}$ with respect to the measure m_{d-1} discussed above. We first make the following simple observation:

1. If f is continuous and has compact support, then f is of course integrable on *every* plane $\mathcal{P}_{t,\gamma}$, and so $\mathcal{R}(f)(t,\gamma)$ is defined for all $(t,\gamma) \in \mathbb{R} \times S^{d-1}$. Moreover it is a continuous function of the pair (t,γ) and has compact support in the t-variable.

2. If f is merely Lebesgue integrable, then f may fail to be measurable or integrable on $\mathcal{P}_{t,\gamma}$ for some (t,γ), and thus $\mathcal{R}(f)(t,\gamma)$ is not defined for those (t,γ).

3. Suppose f is the characteristic function of the set E, that is, $f = \chi_E$. Then $\mathcal{R}(f)(t,\gamma) = m_{d-1}(E_{t,\gamma})$ if $E_{t,\gamma}$ is measurable.

It is this last property that links the Radon transform to our problem. Key estimates in this conclusion involve a maximal "Radon transform" defined by

$$\mathcal{R}^*(f)(\gamma) = \sup_{t\in\mathbb{R}} |\mathcal{R}(f)(t,\gamma)|,$$

as well as corresponding expressions controlling the Lipschitz character of $\mathcal{R}(f)(t,\gamma)$ as a function of t. A basic fact inherent in our analysis is that the regularity of the Radon transform actually improves as the dimension of the underlying space increases.

Theorem 4.5 *Suppose f is continuous and has compact support in $\mathbb{R}^d$ with $d \geq 3$. Then*

$$\int_{S^{d-1}} \mathcal{R}^*(f)(\gamma)\, d\sigma(\gamma) \leq c\left[\|f\|_{L^1(\mathbb{R}^d)} + \|f\|_{L^2(\mathbb{R}^d)}\right] \tag{2}$$

for some constant $c > 0$ that does not depend on f.

An inequality of this type is a typical "a priori" estimate. It is obtained first under some regularity assumption on the function f, and then a limiting argument allows one to pass to the more general case when f belongs to $L^1 \cap L^2$.

We make some comments about the appearance of both the L^1-norm and L^2-norm in (2). The L^2-norm imposes a crucial local control of the kind that is necessary for the desired regularity. (See Exercise 27.) However, without some restriction on f of a global nature, the function f might fail to be integrable on *any* plane $\mathcal{P}_{t,\gamma}$, as the example $f(x) = 1/(1+|x|^{d-1})$ shows. Note that this function belongs to $L^2(\mathbb{R}^d)$ if $d \geq 3$, but not to $L^1(\mathbb{R}^d)$.

The proof of Theorem 4.5 actually gives an essentially stronger result, which we state as a corollary.

Corollary 4.6 *Suppose f is continuous and has compact support in $\mathbb{R}^d$, $d \geq 3$. Then for any α, $0 < \alpha < 1/2$, the inequality* (2) *holds with $\mathcal{R}^*(f)(\gamma)$ replaced by*

$$\sup_{t_1 \neq t_2} \frac{|\mathcal{R}(f)(t_1,\gamma) - \mathcal{R}(f)(t_2,\gamma)|}{|t_1 - t_2|^\alpha}. \tag{3}$$

The proof of the theorem relies on the interplay between the Radon transform and the Fourier transform.

For fixed $\gamma \in S^{d-1}$, we let $\hat{\mathcal{R}}(f)(\lambda,\gamma)$ denote the Fourier transform of $\mathcal{R}(f)(t,\gamma)$ in the t-variable

$$\hat{\mathcal{R}}(f)(\lambda,\gamma) = \int_{-\infty}^{\infty} \mathcal{R}(f)(t,\gamma)e^{-2\pi i\lambda t}\,dt.$$

In particular, we use $\lambda \in \mathbb{R}$ to denote the dual variable of t.

We also write $\hat{f}$ for the Fourier transform of f as a function on $\mathbb{R}^d$, namely

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}\,dx.$$

Lemma 4.7 *If f is continuous with compact support, then for every $\gamma \in S^{d-1}$ we have*

$$\hat{\mathcal{R}}(f)(\lambda,\gamma) = \hat{f}(\lambda\gamma).$$

The right-hand side is just the Fourier transform of f evaluated at the point $\lambda\gamma$.

Proof. For each unit vector γ we use the adapted coordinate system described above: $x = (x_1, \ldots, x_d)$ where γ coincides with the x_d direction. We can then write each $x \in \mathbb{R}^d$ as $x = (u, t)$ with $u \in \mathbb{R}^{d-1}$, $t \in \mathbb{R}$, where $x \cdot \gamma = t = x_d$ and $u = (x_1, \ldots, x_{d-1})$. Moreover

$$\int_{\mathcal{P}_{t,\gamma}} f = \int_{\mathbb{R}^{d-1}} f(u,t)\, du,$$

and Fubini's theorem shows that $\int_{\mathbb{R}^d} f(x)\, dx = \int_{-\infty}^{\infty} \left(\int_{\mathcal{P}_{t,\gamma}} f \right) dt$. Applying this to $f(x)e^{-2\pi i x\cdot(\lambda\gamma)}$ in place of $f(x)$ gives

$$\begin{aligned}\hat{f}(\lambda\gamma) &= \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot(\lambda\gamma)}\, dx = \int_{-\infty}^{\infty} \left(\int_{\mathbb{R}^{d-1}} f(u,t)\, du \right) e^{-2\pi i\lambda t}\, dt \\ &= \int_{-\infty}^{\infty} \left(\int_{\mathcal{P}_{t,\gamma}} f \right) e^{-2\pi i\lambda t}\, dt.\end{aligned}$$

Therefore $\hat{f}(\lambda\gamma) = \hat{\mathcal{R}}(f)(\lambda, \gamma)$, and the lemma is proved.

Lemma 4.8 *If f is continuous with compact support, then*

$$\int_{S^{d-1}} \left(\int_{-\infty}^{\infty} |\hat{\mathcal{R}}(f)(\lambda,\gamma)|^2 |\lambda|^{d-1} d\lambda \right) d\sigma(\gamma) = 2\int_{\mathbb{R}^d} |f(x)|^2 dx.$$

Let us observe the crucial point that the greater the dimension d, the larger the factor $|\lambda|^{d-1}$ as $|\lambda|$ tends to infinity. Hence the greater the dimension, the better the decay of the Fourier transform $\hat{\mathcal{R}}(f)(\lambda, \gamma)$, and so the better the regularity of the Radon transform $\mathcal{R}(f)(t, \gamma)$ as a function of t.

Proof. The Plancherel formula in Chapter 5 guarantees that

$$2\int_{\mathbb{R}^d} |f(x)|^2\, dx = 2\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\, d\xi.$$

Changing to polar coordinates $\xi = \lambda\gamma$ where $\lambda > 0$ and $\gamma \in S^{d-1}$, we obtain

$$2\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\, d\xi = 2\int_{S^{d-1}} \int_0^{\infty} |\hat{f}(\lambda\gamma)|^2 \lambda^{d-1}\, d\lambda\, d\sigma(\gamma).$$

We now observe that a simple change of variables provides

$$\int_{S^{d-1}} \int_0^{\infty} |\hat{f}(\lambda\gamma)|^2 \lambda^{d-1}\, d\lambda\, d\sigma(\gamma) = \int_{S^{d-1}} \int_{-\infty}^{0} |\hat{f}(\lambda\gamma)|^2 |\lambda|^{d-1}\, d\lambda\, d\sigma(\gamma),$$

and the proof is complete once we invoke the result of Lemma 4.7.

The final ingredient in the proof of Theorem 4.5 consists of the following:

Lemma 4.9 *Suppose*

$$F(t) = \int_{-\infty}^{\infty} \hat{F}(\lambda) e^{2\pi i \lambda t}\, d\lambda,$$

where

$$\sup_{\lambda \in \mathbb{R}} |\hat{F}(\lambda)| \leq A \quad \textit{and} \quad \int_{-\infty}^{\infty} |\hat{F}(\lambda)|^2 |\lambda|^{d-1} d\lambda \leq B^2.$$

Then

$$\sup_{t \in \mathbb{R}} |F(t)| \leq c(A+B). \tag{4}$$

Moreover, if $0 < \alpha < 1/2$, *then*

$$|F(t_1) - F(t_2)| \leq c_\alpha |t_1 - t_2|^\alpha (A+B) \quad \textit{for all } t_1,\, t_2. \tag{5}$$

Proof. The first inequality is obtained by considering separately the two cases $|\lambda| \leq 1$ and $|\lambda| > 1$. We write

$$F(t) = \int_{|\lambda| \leq 1} \hat{F}(\lambda) e^{2\pi i \lambda t}\, d\lambda + \int_{|\lambda| > 1} \hat{F}(\lambda) e^{2\pi i \lambda t}\, d\lambda.$$

Clearly, the first integral is bounded by cA. To estimate the second integral it suffices to bound $\int_{|\lambda|>1} |\hat{F}(\lambda)|\, d\lambda$. An application of the Cauchy-Schwarz inequality gives

$$\int_{|\lambda|>1} |\hat{F}(\lambda)| d\lambda \leq \left(\int_{|\lambda|>1} |\hat{F}(\lambda)|^2 |\lambda|^{d-1} d\lambda \right)^{1/2} \left(\int_{|\lambda|>1} |\lambda|^{-d+1} d\lambda \right)^{1/2}.$$

This last integral is convergent precisely when $-d+1 < -1$, which is equivalent to $d > 2$, namely $d \geq 3$, which we assume. Hence $|F(t)| \leq c(A+B)$ as desired.

To establish Lipschitz continuity, we first note that

$$F(t_1) - F(t_2) = \int_{-\infty}^{\infty} \hat{F}(\lambda) \left[e^{2\pi i \lambda t_1} - e^{2\pi i \lambda t_2} \right] d\lambda.$$

Since one has the inequality[2] $|e^{ix} - 1| \leq |x|$, we immediately see that

$$|e^{2\pi i\lambda t_1} - e^{2\pi i\lambda t_2}| \leq c|t_1 - t_2|^\alpha \lambda^\alpha \qquad \text{if } 0 \leq \alpha < 1.$$

We may then write the difference $F(t_1) - F(t_2)$ as a sum of two integrals. The integral over $|\lambda| \leq 1$ is clearly bounded by $cA|t_1 - t_2|^\alpha$. The second integral, the one over $|\lambda| > 1$, can be estimated from above by

$$|t_1 - t_2|^\alpha \int_{|\lambda|>1} |\hat{F}(\lambda)||\lambda|^\alpha \, d\lambda.$$

An application of the Cauchy-Schwarz inequality show that this last integral is majorized by

$$\left(\int_{|\lambda|>1} |\hat{F}(\lambda)|^2|\lambda|^{d-1}\, d\lambda\right)^{1/2} \left(\int_{|\lambda|>1} |\lambda|^{-d+1+2\alpha}\, d\lambda\right)^{1/2} \leq c_\alpha B,$$

since the second integral is finite if $-d + 1 + 2\alpha < -1$, and in particular this holds if $\alpha < 1/2$ when $d \geq 3$. This concludes the proof of the lemma.

We now gather these results to prove the theorem. For each $\gamma \in S^{d-1}$ let

$$F(t) = \mathcal{R}(f)(t, \gamma).$$

Note that with this definition we have

$$\sup_{t\in\mathbb{R}} |F(t)| = \mathcal{R}^*(f)(\gamma).$$

Let

$$A(\gamma) = \sup_\lambda |\hat{F}(\lambda)| \qquad \text{and} \qquad B^2(\gamma) = \int_{-\infty}^{\infty} |\hat{F}(\lambda)|^2|\lambda|^{d-1}\, d\lambda.$$

Then by (4)

$$\sup_{t\in\mathbb{R}} |F(t)| \leq c(A(\gamma) + B(\gamma)).$$

However, we observed that $\hat{F}(\lambda) = \hat{f}(\lambda\gamma)$, and hence

$$A(\gamma) \leq \|f\|_{L^1(\mathbb{R}^d)}.$$

---

[2]The distance in the plane from the point e^{ix} to the point 1 is shorter than the length of the arc on the unit circle joining them.

Therefore,

$$|\mathcal{R}^*(f)(\gamma)|^2 \leq c(A(\gamma)^2 + B(\gamma)^2),$$

and thus

$$\int_{S^{d-1}} |\mathcal{R}^*(f)(\gamma)|^2 \, d\sigma(\gamma) \leq c(\|f\|^2_{L^1(\mathbb{R}^d)} + \|f\|^2_{L^2(\mathbb{R}^d)}),$$

since $\int B^2(\gamma)\, d\sigma(\gamma) = 2\|f\|^2_{L^2}$ by Lemma 4.8. Consequently,

$$\int_{S^{d-1}} \mathcal{R}^*(f)(\gamma) \, d\sigma(\gamma) \leq c(\|f\|_{L^1(\mathbb{R}^d)} + \|f\|_{L^2(\mathbb{R}^d)}).$$

Note that the identity we have used,

$$\mathcal{R}(f)(t,\gamma) = \int_{-\infty}^{\infty} \hat{F}(\lambda) e^{2\pi i \lambda t} \, d\lambda,$$

with $F(t) = \mathcal{R}(f)(t,\gamma)$, is justified for almost every $\gamma \in S^{d-1}$ by the Fourier inversion result in Theorem 4.2 of Chapter 2. Indeed, we have seen that $A(\gamma)$ and $B(\gamma)$ are finite for almost every γ, and thus $\hat{F}$ is integrable for those γ. This completes the proof of the theorem. The corollary follows the same way if we use (5) instead of (4).

We now return to the situation in the plane to see what information we may deduce from the above analysis. The inequality (2) as it stands does not hold when $d = 2$. However, a modification of it does hold, and this will be used in the proof of Theorem 4.4.

If $f \in L^1(\mathbb{R}^d)$ we define

$$\begin{aligned} \mathcal{R}_\delta(f)(t,\gamma) &= \frac{1}{2\delta} \int_{t-\delta}^{t+\delta} \mathcal{R}(f)(s,\gamma)\, ds \\ &= \frac{1}{2\delta} \int_{t-\delta \leq x\cdot\gamma \leq t+\delta} f(x)\, dx. \end{aligned}$$

In this definition of $\mathcal{R}_\delta(f)(t,\gamma)$ we integrate the function f in a small "strip" of width 2δ around the plane $\mathcal{P}_{t,\gamma}$. Thus $\mathcal{R}_\delta$ is an average of Radon transforms.

We let

$$\mathcal{R}^*_\delta(f)(\gamma) = \sup_{t\in\mathbb{R}} |\mathcal{R}_\delta(f)(t,\gamma)|.$$

Theorem 4.10 *If f is continuous with compact support, then*

$$\int_{S^1} \mathcal{R}^*_\delta(f)(\gamma)\, d\sigma(\gamma) \le c(\log 1/\delta)^{1/2} \left(\|f\|_{L^1(\mathbb{R}^2)} + \|f\|_{L^2(\mathbb{R}^2)}\right)$$

when $0 < \delta \le 1/2$.

The same argument as in the proof of Theorem 4.5 applies here, except that we need a modified version of Lemma 4.9. More precisely, let us set

$$F_\delta(t) = \int_{-\infty}^{\infty} \hat{F}(\lambda) \left(\frac{e^{2\pi i(t+\delta)\lambda} - e^{2\pi i(t-\delta)\lambda}}{2\pi i \lambda (2\delta)} \right) d\lambda,$$

and suppose that

$$\sup_\lambda |\hat{F}(\lambda)| \le A \quad \text{and} \quad \int_{-\infty}^{\infty} |\hat{F}(\lambda)|^2 |\lambda|\, d\lambda \le B.$$

Then we claim that

$$\sup_t |F_\delta(t)| \le c(\log 1/\delta)^{1/2}(A+B). \tag{6}$$

Indeed, we use the fact that $|(\sin x)/x| \le 1$ to see that, in the definition of $F_\delta(t)$, the integral over $|\lambda| \le 1$ gives the cA. Also, the integral over $|\lambda| > 1$ can be split and is bounded by the sum

$$\int_{1<|\lambda|\le 1/\delta} |\hat{F}(\lambda)|\, d\lambda + \frac{c}{\delta} \int_{1/\delta \le |\lambda|} |\hat{F}(\lambda)||\lambda|^{-1}\, d\lambda.$$

The first integral above can be estimated by the Cauchy-Schwarz inequality, as follows

$$\begin{aligned}\int_{1<|\lambda|\le 1/\delta} |\hat{F}(\lambda)|\, d\lambda &\le c \left(\int_{1<|\lambda|\le 1/\delta} |\hat{F}(\lambda)|^2 |\lambda|\, d\lambda \right)^{1/2} \left(\int_{1<|\lambda|\le 1/\delta} |\lambda|^{-1}\, d\lambda \right)^{1/2} \\ &\le cB(\log 1/\delta)^{1/2}.\end{aligned}$$

Finally, we also note that

$$\begin{aligned}\frac{c}{\delta} \int_{1/\delta\le|\lambda|} |\hat{F}(\lambda)||\lambda|^{-1}\, d\lambda &\le c \left(\int_{1/\delta\le|\lambda|} |\hat{F}(\lambda)|^2 |\lambda|\, d\lambda \right)^{1/2} \frac{1}{\delta} \left(\int_{1/\delta\le|\lambda|} |\lambda|^{-3}\, d\lambda \right)^{1/2} \\ &\le cB\end{aligned}$$

and this establishes (6), and hence the theorem.

4.2 Regularity of sets when $d \geq 3$

We now extend to the general context the basic estimates for the Radon transform, proved for continuous functions of compact support. This will yield the regularity result formulated in Theorem 4.1.

Proposition 4.11 *Suppose $d \geq 3$, and let f belong to $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Then for a.e. $\gamma \in S^{d-1}$ we can assert the following:*

(a) *f is measurable and integrable on the plane $\mathcal{P}_{t,\gamma}$, for every $t \in \mathbb{R}$.*

(b) *The function $\mathcal{R}(f)(t,\gamma)$ is continuous in t and satisfies a Lipschitz condition with exponent α for each $\alpha < 1/2$. Moreover, the inequality (2) of Theorem 4.5 and its variant with (3) hold for f.*

We prove this in a series of steps.

Step 1. We consider $f = \chi_{\mathcal{O}}$, the characteristic function of a bounded open set $\mathcal{O}$. Here the assertion (a) is evident since $\mathcal{O} \cap \mathcal{P}_{t,\gamma}$ is an open and bounded set in $\mathcal{P}_{t,\gamma}$ and is bounded. Thus $\mathcal{R}(f)(t,\gamma)$ is defined for all (t,γ).

Next we can find a sequence $\{f_n\}$ of non-negative continuous functions of compact support so that for every x, $f_n(x)$ increases to $f(x)$ as $n \to \infty$. Thus $\mathcal{R}(f_n)(t,\gamma) \to \mathcal{R}(f)(t,\gamma)$ for every (t,γ) by the monotone convergence theorem, and also $\mathcal{R}^*(f_n)(\gamma) \to \mathcal{R}^*(f)(\gamma)$ for each $\gamma \in S^{d-1}$. As a result we see that the inequality (2) is valid for $f = \chi_{\mathcal{O}}$, with $\mathcal{O}$ open and bounded.

Step 2. We now consider $f = \chi_E$, where E is a set of measure zero, and take first the case when the set E is bounded. Then we can find a decreasing sequence $\{\mathcal{O}_n\}$ of open and bounded sets, such that $E \subset \mathcal{O}_n$, while $m(\mathcal{O}_n) \to 0$ as $n \to \infty$.

Let $\tilde{E} = \bigcap \mathcal{O}_n$. Since $\tilde{E} \cap \mathcal{P}_{t,\gamma}$ is measurable for every (t,γ), the functions $\mathcal{R}(\chi_{\tilde{E}})(t,\gamma)$ and $\mathcal{R}^*(\chi_{\tilde{E}})(\gamma)$ are well-defined. However, $\mathcal{R}^*(\chi_{\tilde{E}})(\gamma) \leq \mathcal{R}^*(\chi_{\mathcal{O}_n})(\gamma)$, while the $\mathcal{R}^*(\chi_{\mathcal{O}_n})$ decrease. Thus the inequality (2) we have just proved for $f = \chi_{\mathcal{O}_n}$ shows that $\mathcal{R}^*(\chi_{\tilde{E}})(\gamma) = 0$ for a.e. γ. The fact that $E \subset \tilde{E}$ then implies that for a.e. γ, the set $E \cap \mathcal{P}_{t,\gamma}$ has $(d-1)$-dimensional measure zero for every $t \in \mathbb{R}$. This conclusion immediately extends to the case when E is not necessarily bounded, by writing E as a countable union of bounded sets of measure zero. Therefore Corollary 4.2 is proved.

Step 3. Here we assume that f is a bounded measurable function supported on a bounded set. Then by familiar arguments we can find a sequence $\{f_n\}$ of continuous functions that are uniformly bounded,

supported in a fixed compact set, and so that $f_n(x) \to f(x)$ a.e. By the bounded convergence theorem, $\|f_n - f\|_{L^1}$ and $\|f_n - f\|_{L^2}$ both tend to zero as $n \to \infty$, and upon selecting a subsequence if necessary, we can suppose that $\|f_n - f\|_{L^1} + \|f_n - f\|_{L^2} \leq 2^{-n}$. By what we have just proved in Step 2 we have, for a.e. $\gamma \in S^{d-1}$, that $f_n(x) \to f(x)$ on $\mathcal{P}_{t,\gamma}$ a.e. with respect to the measure m_{d-1}, for each $t \in \mathbb{R}$. Thus again by the bounded convergence theorem for those (t, γ), we see that $\mathcal{R}(f_n)(t, \gamma) \to \mathcal{R}(f)(t, \gamma)$, and this limit defines $\mathcal{R}(f)$. Now applying Theorem 4.5 to $f_n - f_{n-1}$ gives

$$\sum_{n=1}^{\infty} \int_{S^{d-1}} \mathcal{R}^*(f_n - f_{n-1})(\gamma)\, d\sigma(\gamma) \leq c \sum_{n=1}^{\infty} 2^{-n} < \infty.$$

This means that

$$\sum_n \sup_t |\mathcal{R}(f_n)(t, \gamma) - \mathcal{R}(f_{n-1})(t, \gamma)| < \infty,$$

for a.e.$\gamma \in S^{d-1}$, and hence for those γ the sequence of functions $\mathcal{R}(f_n)(t, \gamma)$ converges uniformly. As a consequence, for those γ the function $\mathcal{R}(f)(t, \gamma)$ is continuous in t, and the inequality (2) is valid for this f. The inequality with (3) is deduced in the same way.

Finally, we deal with the general f in $L^1 \cap L^2$ by approximating it by a sequence of bounded functions each with bounded support. The details of the argument are similar to the case treated above and are left to the reader.

Observe that the special case $f = \chi_E$ of the proposition gives us Theorem 4.1.

4.3 Besicovitch sets have dimension 2

Here we prove Theorem 4.4, that any Besicovitch set necessarily has Hausdorff dimension 2. We use Theorem 4.10, namely, the inequality

$$\int_{S^1} \mathcal{R}^*_\delta(f)(\gamma)\, d\sigma(\gamma) \leq c(\log 1/\delta)^{1/2} \left(\|f\|_{L^1(\mathbb{R}^2)} + \|f\|_{L^2(\mathbb{R}^2)}\right).$$

This inequality was proved under the assumption that f was continuous and had compact support. In the present situation it goes over without difficulty to the general case where $f \in L^1 \cap L^2$, by an easy limiting argument, since it is clear that $\mathcal{R}^*_\delta(f_n)(\gamma)$ converges to $\mathcal{R}^*_\delta(f)(\gamma)$ for all γ if $f_n \to f$ in the L^1-norm.

Now suppose F is a Besicovitch set and α is fixed with $0 < \alpha < 2$. Assume that $F \subset \bigcup_{i=1}^{\infty} B_i$ is a covering, where B_i are balls with diameter less than a given number. We must show that

$$\sum_i (\text{diam } B_i)^\alpha \geq c_\alpha > 0.$$

We proceed in two steps, considering first a simple situation that will make clear the idea of the proof.

Case 1. We suppose first that all the balls B_i have the same diameter δ (with $\delta \leq 1/2$) and also that there are only a finite number, say N, of balls in the covering. We must prove that $N\delta^\alpha \geq c_\alpha$.

Let B_i^* denote the double of B_i and $F^* = \bigcup_i B_i^*$. Then, we clearly have

$$m(F^*) \leq cN\delta^2.$$

Since F is a Besicovitch set, for each $\gamma \in S^1$ there is a segment s_γ of unit length, perpendicular to γ, and which is contained in F. Also, by construction, any translate by less than δ of a point in s_γ must belong to F^*. Hence

$$\mathcal{R}_\delta^*(\chi_{F^*})(\gamma) \geq 1 \quad \text{for every } \gamma.$$

If we take $f = \chi_{F^*}$ in the inequality (6), and note that the Cauchy-Schwarz inequality implies

$$\|\chi_{F^*}\|_{L^1(\mathbb{R}^2)} \leq c\|\chi_{F^*}\|_{L^2(\mathbb{R}^2)} \leq c(m(F^*))^{1/2},$$

then we obtain

$$c \leq N^{1/2}\delta(\log 1/\delta)^{1/2}.$$

This implies $N\delta^\alpha \geq c$ for $\alpha < 2$.

Case 2. We now treat the general case. Suppose $F \subset \bigcup_{i=1}^{\infty} B_i$, where the balls B_i each have diameter less than 1. For each integer k let N_k be the number of balls in the collection $\{B_i\}$ for which

$$2^{-k-1} \leq \text{diam } B_i \leq 2^{-k}.$$

We need to show that $\sum_{k=0}^{\infty} N_k 2^{-k\alpha} \geq c_\alpha$. In fact, we shall prove the stronger result that there exists a positive integer k' such that $N_{k'}2^{-k'\alpha} \geq c_\alpha$.

Let

$$F_k = F \bigcap \left(\bigcup_{2^{-k-1} \leq \text{diam } B_i \leq 2^{-k}} B_i \right),$$

and let

$$F_k^* = \bigcup_{2^{-k-1} \leq \text{diam } B_i \leq 2^{-k}} B_i^*,$$

where B_i^* denotes the double of B_i. Then we note that

$$m_1(F_k^*) \leq cN_k 2^{-2k} \quad \text{for all } k.$$

Since F is a Besicovitch set, for each $\gamma \in S^1$ there is a segment s_γ of unit length entirely contained in F. We now make precise the fact that for some k, a large proportion of s_γ belongs to F_k.

We pick a sequence of real numbers $\{a_k\}_{k=0}^{\infty}$ such that $0 \leq a_k \leq 1$, $\sum a_k = 1$, but a_k does not tend to zero too quickly. For instance, we may choose $a_k = c_\epsilon 2^{-\epsilon k}$ with $c_\epsilon = 1 - 2^{-\epsilon}$, and $\epsilon > 0$ but ϵ sufficiently small.

Then, for some k we must have

$$m_1(s_\gamma \cap F_k) \geq a_k.$$

Otherwise, since $F = \bigcup F_k$, we would have

$$m_1(s_\gamma \cap F) < \sum a_k = 1,$$

and this contradicts the fact that $m_1(s_\gamma \cap F) = 1$, since s_γ is entirely contained in F.

Therefore, with this k, we must have

$$\mathcal{R}^*_{2^{-k}}(\chi_{F_k^*})(\gamma) \geq a_k,$$

because any point of distance less than 2^{-k} from F_k must belong to F_k^*. Since the choice of k may depend on γ, we let

$$E_k = \{\gamma : \ \mathcal{R}^*_{2^{-k}}(\chi_{F_k^*})(\gamma) \geq a_k\}.$$

By our previous observations, we have

$$S^1 = \bigcup_{k=1}^{\infty} E_k,$$

and so for at least one k, which we denote by k', we have

$$m(E_{k'}) \geq 2\pi a_{k'},$$

for otherwise $m(S_1) < 2\pi \sum a_k = 2\pi$. As a result

$$\begin{aligned} 2\pi a_{k'}^2 &= 2\pi a_{k'} a_{k'} \\ &\leq \int_{E_{k'}} a_{k'}\, d\sigma(\gamma) \\ &\leq \int_{S_1} \mathcal{R}^*_{2^{-k'}}(\chi_{F^*_{k'}})(\gamma)\, d\sigma(\gamma). \end{aligned}$$

By the fundamental inequality (6) we get

$$a_{k'}^2 \leq c(\log 2^{k'})^{1/2} \|\chi_{F^*_{k'}}\|_{L^2(\mathbb{R}^2)}.$$

Recalling that by our choice $a_k \approx 2^{-\epsilon k}$, and noting that $\|\chi_{F^*_{k'}}\|_{L^2} \leq cN_{k'}^{1/2} 2^{-k'}$, we obtain

$$2^{(1-2\epsilon)k'} \leq c(\log 2^{k'})^{1/2} N_{k'}^{1/2}.$$

Finally, this last inequality guarantees that $N_{k'} 2^{-\alpha k'} \geq c_\alpha$ as long as $4\epsilon < 2 - \alpha$.

This concludes the proof of the theorem.

4.4 Construction of a Besicovitch set

There are a number of different constructions of Besicovitch sets. The one we have chosen to describe here involves the concept of self-replicating sets, an idea that permeates much of the discussion of this chapter.

We consider the Cantor set of constant dissection $\mathcal{C}_{1/2}$, which for simplicity we shall write as $\mathcal{C}$, and which is defined in Exercise 3, Chapter 1. Note that $\mathcal{C} = \bigcap_{k=0}^{\infty} C_k$, where $C_0 = [0, 1]$, and C_k is the union of 2^k closed intervals of length 4^{-k} obtained by removing from C_{k-1} the 2^{k-1} centrally situated open intervals of length $\frac{1}{2} \cdot 4^{-k+1}$. The set $\mathcal{C}$ can also be represented as the set of points $x \in [0, 1]$ of the form $x = \sum_{k=1}^{\infty} \epsilon_k / 4^k$, with ϵ_k either 0 or 3.

We now place a copy of $\mathcal{C}$ on the x-axis of the plane $\mathbb{R}^2 = \{(x, y)\}$, and a copy of $\frac{1}{2}\mathcal{C}$ on the line $y = 1$. That is, we put $E_0 = \{(x, y) : x \in \mathcal{C},\ y = 0\}$ and $E_1 = \{(x, y) : 2x \in \mathcal{C},\ y = 1\}$. The set F that will play the central role is defined as the union of all line segments that join a point of E_0 with a point of E_1. (See Figure 13.)

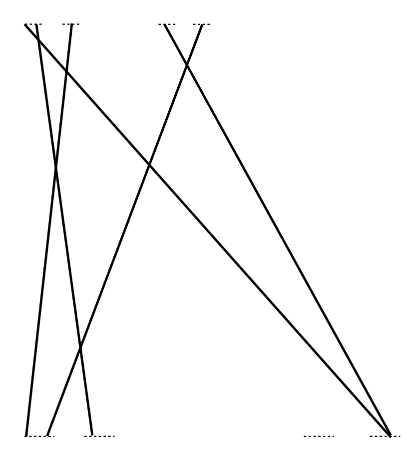


Figure 13. Several line segments joining E_0 with E_1

Theorem 4.12 *The set F is compact and of two-dimensional measure zero. It contains a translate of any unit line segment whose slope is a number s that lies outside the intervals* $(-1, 2)$.

Once the theorem is proved, our job is done. Indeed, a finite union of rotations of the set F contains unit segments of any slope, and that set is therefore a Besicovitch set.

The proof of the required properties of the set F amounts to showing the following paradoxical facts about the set $\mathcal{C} + \lambda\mathcal{C}$, for $\lambda > 0$. Here $\mathcal{C} + \lambda\mathcal{C} = \{x_1 + \lambda x_2 : \ x_1 \in \mathcal{C}, x_2 \in \mathcal{C}\}$:

- $\mathcal{C} + \lambda\mathcal{C}$ has one-dimensional measure zero, for a.e. λ.
- $\mathcal{C} + \frac{1}{2}\mathcal{C}$ is the interval $[0, 3/2]$.

Let us see how these two assertions imply the theorem. First, we note that the set F is closed (and hence compact), because both E_0 and E_1 are closed. Next observe that with $0 < y < 1$, the slice F^y of the set F is exactly $(1-y)\mathcal{C} + \frac{y}{2}\mathcal{C}$. This set is obtained from the set $\mathcal{C} + \lambda\mathcal{C}$, where $\lambda = y/(2(1-y))$, by scaling with the factor $1 - y$. Hence F^y is of measure zero whenever $\mathcal{C} + \lambda\mathcal{C}$ is also of measure zero. Moreover, under the mapping $y \mapsto \lambda$, sets of measure zero in $(0, \infty)$ correspond to sets of measure zero in $(0, 1)$. (For this see, for example, Exercise 8 in Chapter 1, or Problem 1 in Chapter 6.) Therefore, the first assertion and Fubini's theorem prove that the (two-dimensional) measure of F is zero.

Finally the slope s of the segment joining the point $(x_0, 0)$, with the point $(x_1, 1)$ is $s = 1/(x_1 - x_0)$. Thus the quantity s can be realized if

$x_1 \in \mathcal{C}/2$ and $x_0 \in \mathcal{C}$, that is, if $1/s \in \mathcal{C}/2 - \mathcal{C}$. However, by an obvious symmetry $\mathcal{C} = 1 - \mathcal{C}$, and so the condition becomes $1/s \in \mathcal{C}/2 + \mathcal{C} - 1$, which by the second assertion is $1/s \in [-1, 1/2]$. This last is equivalent with $s \notin (-1, 2)$.

Our task therefore remains the proof of the two assertions above. The proof of the second is nearly trivial. In fact,

$$\frac{2}{3}\left(\mathcal{C} + \frac{1}{2}\mathcal{C}\right) = \frac{2}{3}\mathcal{C} + \frac{1}{3}\mathcal{C},$$

and this set consists of all x of the form $x = \sum_{k=1}^{\infty}\left(\frac{2\epsilon_k}{3} + \frac{\epsilon_k'}{3}\right)4^{-k}$, where ϵ_k and ϵ_k' are independently 0 or 3. Since then $\frac{2\epsilon_k}{3} + \frac{\epsilon_k'}{3}$ can take any of the values 0, 1, 2, or 3, we have that $\frac{2}{3}\left(\mathcal{C} + \frac{1}{2}\mathcal{C}\right) = [0, 1]$, and hence $\mathcal{C} + \frac{1}{2}\mathcal{C} = [0, 3/2]$.

The proof that $m(\mathcal{C} + \lambda\mathcal{C}) = 0$ for a.e. λ

We come to the main point: that $\mathcal{C} + \lambda\mathcal{C}$ has measure zero for almost all λ. We show this by examining the self-replicating properties of the sets $\mathcal{C}$ and $\mathcal{C} + \lambda\mathcal{C}$.

We know that $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$, where $\mathcal{C}_1$ and $\mathcal{C}_2$ are two similar copies of $\mathcal{C}$, obtained with similarity ratio 1/4, and given by $\mathcal{C}_1 = \frac{1}{4}\mathcal{C}$ and $\mathcal{C}_2 = \frac{1}{4}\mathcal{C} + \frac{3}{4}$. Thus $\mathcal{C}_1 \subset [0, 1/4]$ and $\mathcal{C}_2 \subset [3/4, 1]$. Iterating ℓ times this decomposition of $\mathcal{C}$, that is, reaching the ℓth "generation," we can write

$$\mathcal{C} = \bigcup_{1 \le j \le 2^\ell} \mathcal{C}_j^\ell, \tag{7}$$

with $\mathcal{C}_1^\ell = (1/4)^\ell\mathcal{C}$ and each $\mathcal{C}_j^\ell$ a translate of $\mathcal{C}_1^\ell$.

We consider in the same way the set

$$\mathcal{K}(\lambda) = \mathcal{C} + \lambda\mathcal{C},$$

and we shall sometimes omit the λ and write $\mathcal{K}(\lambda) = \mathcal{K}$, when this causes no confusion. By its definition we have

$$\mathcal{K} = \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3 \cup \mathcal{K}_4,$$

where $\mathcal{K}_1 = \mathcal{C}_1 + \lambda\mathcal{C}_1$, $\mathcal{K}_2 = \mathcal{C}_1 + \lambda\mathcal{C}_2$, $\mathcal{K}_3 = \mathcal{C}_2 + \lambda\mathcal{C}_1$, and $\mathcal{K}_4 = \mathcal{C}_2 + \lambda\mathcal{C}_2$. An iteration of this decomposition using (7) gives

$$\mathcal{K} = \bigcup_{1 \le i \le 4^\ell} \mathcal{K}_i^\ell, \tag{8}$$

where each $\mathcal{K}_i^\ell$ equals $\mathcal{C}_{j_1}^\ell + \lambda\mathcal{C}_{j_2}^\ell$ for a pair of indices j_1, j_2. In fact, this relation among the indices sets up a bijection between the i with $1 \le i \le 4^\ell$, and the pair j_1, j_2 with $1 \le j_1 \le 2^\ell$ and $1 \le j_2 \le 2^\ell$. Note that each $\mathcal{K}_i^\ell$ is a translate of $\mathcal{K}_1^\ell$, and each $\mathcal{K}_i^\ell$ is also obtained from $\mathcal{K}$ by a similarity of ratio $4^{-\ell}$. Now note that $\mathcal{C} = \mathcal{C}/4 \bigcup(\mathcal{C}/4 + 3/4)$ implies that

$$\begin{aligned}\mathcal{K}(\lambda) = \mathcal{C} + \lambda\mathcal{C} &= (\mathcal{C} + \frac{\lambda}{4}\mathcal{C}) \cup (\mathcal{C} + \frac{\lambda}{4}\mathcal{C} + \frac{3\lambda}{4}) \\ &= \mathcal{K}(\lambda/4) \cup (\mathcal{K}(\lambda/4) + \frac{3\lambda}{4}).\end{aligned}$$

Thus $\mathcal{K}(\lambda)$ has measure zero if and only if $\mathcal{K}(\lambda/4)$ has measure zero. Hence it suffices to prove that $\mathcal{K}(\lambda)$ has measure zero for a.e. $\lambda \in [1, 4]$.

After these preliminaries let us observe that we immediately obtain that $m(\mathcal{K}(\lambda)) = 0$ for some special λ's, those for which the following **coincidence** takes place: for some ℓ and a pair i and i' with $i \neq i'$,

$$\mathcal{K}_i^\ell(\lambda) = \mathcal{K}_{i'}^\ell(\lambda).$$

Indeed, if we have this coincidence, then (8) gives

$$m(\mathcal{K}(\lambda)) \le \sum_{i=1,\ i\neq i'}^{4^\ell} m(\mathcal{K}_i^\ell(\lambda)) = (4^\ell - 1)4^{-\ell} m(\mathcal{K}(\lambda)),$$

and this implies $m(\mathcal{K}(\lambda)) = 0$.

The key insight below is that, in a quantitative sense, the λ's for which this coincidence takes place are "dense" relative to the size of ℓ. More precisely, we have the following.

Proposition 4.13 *Suppose λ_0 and ℓ are given, with $1 \le \lambda_0 \le 4$ and ℓ a positive integer. Then, there exist a $\overline{\lambda}$ and a pair i, i' with $i \neq i'$ such that*

$$\mathcal{K}_i^\ell(\overline{\lambda}) = \mathcal{K}_{i'}^\ell(\overline{\lambda}) \quad \textit{and} \quad |\overline{\lambda} - \lambda_0| \le c4^{-\ell}. \tag{9}$$

Here c is a constant independent of λ_0 and ℓ.

This is proved on the basis of the following observation.

Lemma 4.14 *For every λ_0 there is a pair $1 \le i_1, i_2 \le 4$, with $i_1 \neq i_2$ such that $\mathcal{K}_{i_1}(\lambda_0)$ and $\mathcal{K}_{i_2}(\lambda_0)$ intersect.*

Proof. Indeed, if the $\mathcal{K}_i$ are disjoint for $1 \leq i \leq 4$ then for sufficiently small δ the $\mathcal{K}_i^\delta$ are also disjoint. Here we have used the notation that F^δ denotes the set of points of distance less than δ from F. (See Lemma 3.1 in Chapter 1.) However, $\mathcal{K}^\delta = \bigcup_{i=1}^4 \mathcal{K}_i^\delta$, and by similarity $m(\mathcal{K}^{4\delta}) = 4m(\mathcal{K}_i^\delta)$. Thus by the disjointness of the $\mathcal{K}_i^\delta$ we have $m(\mathcal{K}^\delta) = m(\mathcal{K}^{4\delta})$, which is a contradiction, since $\mathcal{K}^{4\delta} - \mathcal{K}^\delta$ contains an open ball (of radius $3\delta/2$). The lemma is therefore proved.

Now apply the lemma for our given λ_0 and write $\mathcal{K}_{i_1} = \mathcal{C}_{\mu_1} + \lambda_0\mathcal{C}_{\nu_1}$, $\mathcal{K}_{i_2} = \mathcal{C}_{\mu_2} + \lambda_0\mathcal{C}_{\nu_2}$, where the μ's and ν's are either 1 or 2. However, since $i_1 \neq i_2$ we have $\mu_1 \neq \mu_2$ or $\nu_1 \neq \nu_2$ (or both). Assume for the moment that $\nu_1 \neq \nu_2$.

The fact that $\mathcal{K}_{i_1}(\lambda_0)$ and $\mathcal{K}_{i_2}(\lambda_0)$ intersect means that there are pairs of numbers (a, b) and (a', b'), with $a \in \mathcal{C}_{\mu_1}$, $b \in \mathcal{C}_{\nu_1}$, $a' \in \mathcal{C}_{\mu_2}$, and $b' \in \mathcal{C}_{\nu_2}$ such that

$$a + \lambda_0 b = a' + \lambda_0 b'. \tag{10}$$

Note that the fact that $\nu_1 \neq \nu_2$ means that $|b - b'| \geq 1/2$. Next, looking at the ℓth generation we find via (7) that there are indices $1 \leq j_1, j_2, j_1', j_2' \leq 2^\ell$, so that $a \in \mathcal{C}_{j_1}^\ell \subset \mathcal{C}_{\mu_1}$, $b \in \mathcal{C}_{j_2}^\ell \subset \mathcal{C}_{\nu_1}$, $a' \in \mathcal{C}_{j_1'}^\ell \subset \mathcal{C}_{\mu_2}$, $b' \in \mathcal{C}_{j_2'}^\ell \subset \mathcal{C}_{\nu_2}$. We also observe that the above sets are translates of each other, that is, $\mathcal{C}_{j_1}^\ell = \mathcal{C}_{j_1'}^\ell + \tau_1$ and $\mathcal{C}_{j_2}^\ell = \mathcal{C}_{j_2'}^\ell + \tau_2$, with $|\tau_k| \leq 1$. Hence if i and i' correspond to the pairs (j_1, j_2) and (j_1', j_2'), respectively, we have

$$\mathcal{K}_i^\ell(\lambda) = \mathcal{K}_{i'}^\ell(\lambda) + \tau(\lambda) \quad \text{with } \tau(\lambda) = \tau_1 + \lambda\tau_2. \tag{11}$$

Now let (A, B) be the pair that corresponds to (a', b') under the above translations, namely

$$A = a' + \tau_1, \qquad B = b' + \tau_2. \tag{12}$$

We claim there is a $\overline{\lambda}$ such that

$$A + \overline{\lambda} B = a' + \overline{\lambda} b'. \tag{13}$$

In fact, by (12) we have put B in $\mathcal{C}_{j_2}^\ell \subset \mathcal{C}_{\nu_1}$, while b' is in $\mathcal{C}_{j_2'}^\ell \subset \mathcal{C}_{\nu_2}$. Thus $|B - b'| \geq 1/2$, since $\nu_1 \neq \nu_2$. We can therefore solve (13) by taking $\overline{\lambda} = (A - a')/(b' - B)$. Now we compare this with (10), and get $\lambda_0 = (a - a')/(b' - b)$. Moreover, $|A - a| \leq 4^{-\ell}$ and $|B - b| \leq 4^{-\ell}$, since A and a both lie in $\mathcal{C}_{j_1}^\ell$, and B and b lie in $\mathcal{C}_{j_2}^\ell$. This yields the inequality

$$|\overline{\lambda} - \lambda_0| \leq c4^{-\ell}. \tag{14}$$

Also, (12) and (13) clearly imply $\tau(\overline{\lambda}) = \tau_1 + \overline{\lambda}\tau_2 = 0$, and this together with (11) proves the coincidence.

Therefore our proposition is proved under the restriction we made earlier that $\nu_1 \neq \nu_2$. The situation where instead $\mu_1 \neq \mu_2$ is obtained from the case $\nu_1 \neq \nu_2$ if we replace λ_0 by λ_0^{-1}. Note that $\mathcal{K}_i^\ell(\lambda_0) = \mathcal{K}_{i'}^\ell(\lambda_0)$ if and only if $\mathcal{C}_{j_1}^\ell + \lambda_0 \mathcal{C}_{j_2}^\ell = \mathcal{C}_{j_1'}^\ell + \lambda_0 \mathcal{C}_{j_2'}^\ell$ and this is the same as $\mathcal{C}_{j_2}^\ell + \lambda_0^{-1}\mathcal{C}_{j_1}^\ell = \mathcal{C}_{j_2'}^\ell + \lambda_0^{-1}\mathcal{C}_{j_1'}^\ell$. This allows us to reduce to the case $\mu_1 \neq \mu_2$, since $\mathcal{C}_{j_1}^\ell \subset \mathcal{C}_{\mu_1}$ and $\mathcal{C}_{j_1'}^\ell \subset \mathcal{C}_{\mu_2}$. Here the fact that $1 \leq \lambda_0 \leq 4$ gives $\lambda_0^{-1} \leq 1$ and guarantees that the constant c in (9) can be taken to be independent of λ_0. The proposition is therefore established.

Note that as a consequence, the following holds near the points $\overline{\lambda}$ where the coincidence (9) takes place: If $|\lambda - \overline{\lambda}| \leq \epsilon 4^{-\ell}$, then

$$\mathcal{K}_i^\ell(\lambda) = \mathcal{K}_{i'}^\ell(\lambda) + \tau(\lambda) \quad \text{with } |\tau(\lambda)| \leq \epsilon 4^{-\ell}. \tag{15}$$

In fact, this is (11) together with the observation that

$$|\tau(\lambda)| = |\tau(\lambda) - \tau(\overline{\lambda})| \leq |\lambda - \overline{\lambda}|,$$

since $|\tau(\lambda)| = \tau_1 + \lambda\tau_2$ and $|\tau_2| \leq 1$.

The assertion (15) leads to the following more elaborate version of itself:

> *There is a set Λ of full measure such that whenever $\lambda \in \Lambda$ and $\epsilon > 0$ are given, there are ℓ and a pair i, i' so that* (15) *holds.*[3]

Indeed, for fixed $\epsilon > 0$, let Λ_ϵ denote the set of λ that satisfies (15) for some ℓ, i and i'. For any interval I of length not exceeding 1, we have

$$m(\Lambda_\epsilon \cap I) \geq \epsilon 4^{-\ell} \geq c^{-1}\epsilon m(I),$$

because of (9) and (15). Thus Λ_ϵ^c has no points of Lebesgue density, hence Λ_ϵ^c has measure zero, and thus Λ_ϵ is a set of full measure. (See Corollary 1.5 in Chapter 3.) Since $\Lambda = \bigcap_\epsilon \Lambda_\epsilon$, and Λ_ϵ decreases with ϵ, we see that Λ also has full measure and our assertion is proved.

Finally, our theorem will be established once we show that $m(\mathcal{K}(\lambda)) = 0$ whenever $\lambda \in \Lambda$. To prove this, we assume contrariwise that $m(\mathcal{K}(\lambda)) > 0$. Using again the point of density argument, there must be for any

---

[3]The terminology that Λ has "full measure" means that its complement has measure zero.

$0 < \delta < 1$, a non-empty open interval I with $m(\mathcal{K}(\lambda) \cap I) \geq \delta m(I)$. We then fix δ with $1/2 < \delta < 1$ and proceed. With this fixed δ, we select ϵ used below as $\epsilon = m(I)(1 - \delta)$. Next, find ℓ, i, and i' for which (15) holds. The existence of such indices is guaranteed by the hypothesis that $\lambda \in \Lambda$.

We then consider the two similarities (of ratio $4^{-\ell}$) that map $\mathcal{K}(\lambda)$ to $\mathcal{K}_i^\ell(\lambda)$ and $\mathcal{K}_{i'}^\ell(\lambda)$, respectively. These take the interval I to corresponding intervals I_i and $I_{i'}$, respectively, with $m(I_i) = m(I_{i'}) = 4^{-\ell} m(I)$. Moreover,

$$m(\mathcal{K}_i^\ell \cap I_i) \geq \delta m(I_i) \quad \text{and} \quad m(\mathcal{K}_{i'}^\ell \cap I_{i'}) \geq \delta m(I_{i'}).$$

Also, as in (15), $I_{i'} = I_i + \tau(\lambda)$, with $|\tau(\lambda)| \leq \epsilon 4^{-\ell}$. This shows that

$$m(I_i \cap I_{i'}) \geq m(I_i) - \tau(\lambda) \geq 4^{-\ell} m(I) - \epsilon 4^{-\ell} \geq \delta m(I_i),$$

since $\epsilon 4^{-\ell} = (1 - \delta) m(I_i)$. Thus $m(I_i - I_i \cap I_{i'}) \leq (1 - \delta) m(I_i)$, and

$$\begin{aligned} m(\mathcal{K}_i^\ell \cap I_i \cap I_{i'}) &\geq m(\mathcal{K}_i^\ell \cap I_i) - m(I_i - I_i \cap I_{i'}) \\ &\geq (2\delta - 1) m(I_i) \\ &> \frac{1}{2} m(I_i) \geq \frac{1}{2} m(I_i \cap I_{i'}). \end{aligned}$$

So $m(\mathcal{K}_i^\ell \cap I_i \cap I_{i'}) > \frac{1}{2} m(I_i \cap I_{i'})$ and the same holds for i' in place of i. Hence $m(\mathcal{K}_i^\ell \cap \mathcal{K}_{i'}^\ell) > 0$, and this contradicts the decomposition (8) and the fact that $m(\mathcal{K}_i^\ell) = 4^{-\ell} m(\mathcal{K})$ for every i. Therefore we obtain that $m(\mathcal{K}(\lambda)) = 0$ for every $\lambda \in \Lambda$, and the proof of Theorem 4.12 is now complete.

5 Exercises

1. Show that the measure m_α is not σ-finite on $\mathbb{R}^d$ if $\alpha < d$.

2. Suppose E_1 and E_2 are two compact subsets of $\mathbb{R}^d$ such that $E_1 \cap E_2$ contains at most one point. Show directly from the definition of the exterior measure that if $0 < \alpha \leq d$, and $E = E_1 \cup E_2$, then

$$m_\alpha^*(E) = m_\alpha^*(E_1) + m_\alpha^*(E_2).$$

[Hint: Suppose $E_1 \cap E_2 = \{x\}$, let B_ϵ denote the open ball centered at x and of diameter ϵ, and let $E^\epsilon = E \cap B_\epsilon^c$. Show that

$$m_\alpha^*(E^\epsilon) \geq \mathcal{H}_\alpha^\epsilon(E) \geq m_\alpha^*(E) - \mu(\epsilon) - \epsilon^\alpha,$$

where $\mu(\epsilon) \to 0$. Hence $m_\alpha^*(E^\epsilon) \to m_\alpha^*(E)$.]

3. Prove that if $f : [0,1] \to \mathbb{R}$ satisfies a Lipschitz condition of exponent $\gamma > 1$, then f is a constant.

4. Suppose $f : [0,1] \to [0,1] \times [0,1]$ is surjective and satisfies a Lipschitz condition

$$|f(x) - f(y)| \leq C|x - y|^\gamma.$$

Prove that $\gamma \leq 1/2$ directly, without using Theorem 2.2.

[Hint: Divide $[0,1]$ into N intervals of equal length. The image of each sub-interval is contained in a ball of volume $O(N^{-2\gamma})$, and the union of all these balls must cover the square.]

5. Let $f(x) = x^k$ be defined on $\mathbb{R}$, where k is a positive integer and let E be a Borel subset of $\mathbb{R}$.

(a) Show that if $m_\alpha(E) = 0$ for some α, then $m_\alpha(f(E)) = 0$.

(b) Prove that $\dim(E) = \dim f(E)$.

6. Let $\{E_k\}$ be a sequence of Borel sets in $\mathbb{R}^d$. Show that if $\dim E_k \leq \alpha$ for some α and all k, then

$$\dim \bigcup_k E_k \leq \alpha.$$

7. Prove that the $(\log 2/\log 3)$-Hausdorff measure of the Cantor set is precisely equal to 1.

[Hint: Suppose we have a covering of $\mathcal{C}$ by finitely many closed intervals $\{I_j\}$. Then there exists another covering of $\mathcal{C}$ by intervals $\{I'_\ell\}$ each of length 3^{-k} for some k, such that $\sum_j |I_j|^\alpha \geq \sum_\ell |I'_\ell|^\alpha \geq 1$, where $\alpha = \log 2/\log 3$.]

8. Show that the Cantor set of constant dissection, $\mathcal{C}_\xi$, in Exercise 3 of Chapter 1 has strict Hausdorff dimension $\log 2/\log(2/(1-\xi))$.

9. Consider the set $\mathcal{C}_{\xi_1} \times \mathcal{C}_{\xi_2}$ in $\mathbb{R}^2$, with $\mathcal{C}_\xi$ as in the previous exercise. Show that $\mathcal{C}_{\xi_1} \times \mathcal{C}_{\xi_2}$ has strict Hausdorff dimension $\dim(\mathcal{C}_{\xi_1}) + \dim(\mathcal{C}_{\xi_2})$.

10. Construct a Cantor-like set (as in Exercise 4, Chapter 1) that has Lebesgue measure zero, yet Hausdorff dimension 1.

[Hint: Choose $\ell_1, \ell_2, \ldots, \ell_k, \ldots$ so that $1 - \sum_{j=1}^k 2^{j-1}\ell_j$ tends to zero sufficiently slowly as $k \to \infty$.]

11. Let $\mathcal{D} = \mathcal{D}_\mu$ be the Cantor dust in $\mathbb{R}^2$ given as the product $\mathcal{C}_\xi \times \mathcal{C}_\xi$, with $\mu = (1-\xi)/2$.

(a) Show that for any real number λ, the set $\mathcal{C}_\xi + \lambda\mathcal{C}_\xi$ is similar to the projection of $\mathcal{D}$ on the line in $\mathbb{R}^2$ with slope $\lambda = \tan\theta$.

(b) Note that among the Cantor sets $\mathcal{C}_\xi$, the value $\xi = 1/2$ is critical in the construction of the Besicovitch set in Section 4.4. In fact, prove that with $\xi > 1/2$, then $\mathcal{C}_\xi + \lambda\mathcal{C}_\xi$ has Lebesgue measure zero for *every* λ. See also Problem 10 below.

[Hint: $m_\alpha(\mathcal{C}_\xi + \lambda\mathcal{C}_\xi) < \infty$ for $\alpha = \dim\mathcal{D}_\mu$.]

12. Define a primitive one-dimensional "measure" $\tilde{m}_1$ as

$$\tilde{m}_1 = \inf\sum_{k=1}^{\infty} \text{diam } F_k, \qquad E \subset \bigcup_{k=1}^{\infty} F_k.$$

This is akin to the one-dimensional exterior measure m_α^*, $\alpha = 1$, except that no restriction is placed on the size of the diameters F_k.

Suppose I_1 and I_2 are two *disjoint* unit segments in $\mathbb{R}^d$, $d \geq 2$, with $I_1 = I_2 + h$, and $|h| < \epsilon$. Then observe that $\tilde{m}_1(I_1) = \tilde{m}_1(I_2) = 1$, while $\tilde{m}_1(I_1 \cup I_2) \leq 1 + \epsilon$. Thus

$$\tilde{m}_1(I_1 \cup I_2) < \tilde{m}_1(I_1) + \tilde{m}_1(I_2) \qquad \text{when } \epsilon < 1;$$

hence $\tilde{m}_1$ fails to be additive.

13. Consider the von Koch curve $\mathcal{K}^\ell$, $1/4 < \ell < 1/2$, as defined in Section 2.1. Prove for it the analogue of Theorem 2.7: the function $t \mapsto \mathcal{K}^\ell(t)$ satisfies a Lipschitz condition of exponent $\gamma = \log(1/\ell)/\log 4$. Moreover, show that the set $\mathcal{K}^\ell$ has strict Hausdorff dimension $\alpha = 1/\gamma$.

[Hint: Show that if $\mathcal{O}$ is the shaded open triangle indicated in Figure 14, then $\mathcal{O} \supset S_0(\mathcal{O}) \cup S_1(\mathcal{O}) \cup S_2(\mathcal{O}) \cup S_3(\mathcal{O})$, where $S_0(x) = \ell x$, $S_1(x) = \rho_\theta(\ell x) + a$, $S_2(x) = \rho_\theta^{-1}(\ell x) + c$, and $S_3(x) = \ell x + b$, with ρ_θ the rotation of angle θ. Note that the sets $S_j(\mathcal{O})$ are disjoint.]

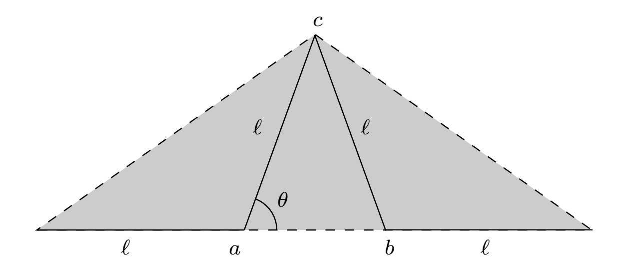


Figure 14. The open set $\mathcal{O}$ in Exercise 13

14. Show that if $\ell < 1/2$, the von Koch curve $t \mapsto \mathcal{K}^\ell(t)$ in Exercise 13 is a simple curve.

[Hint: Observe that if $t = \sum_{j=1}^{\infty} a_j/4^j$, with $a_j = 0, 1, 2$, or 3, then

$$\{\mathcal{K}(t)\} = \bigcap_{j=1}^{\infty} S_{a_j}\left(\cdots S_{a_2}\left(S_{a_1}(\overline{\mathcal{O}})\right)\right).]$$

15. Note that if we take $\ell = 1/2$ in the definition of the von Koch curve in Exercise 13 we get a "space-filling" curve, one that fills the right triangle whose vertices are $(0,0)$, $(1,0)$, and $(1/2,1/2)$. The first three steps of the construction are as in Figure 15, with the intervals traced out in the indicated order.

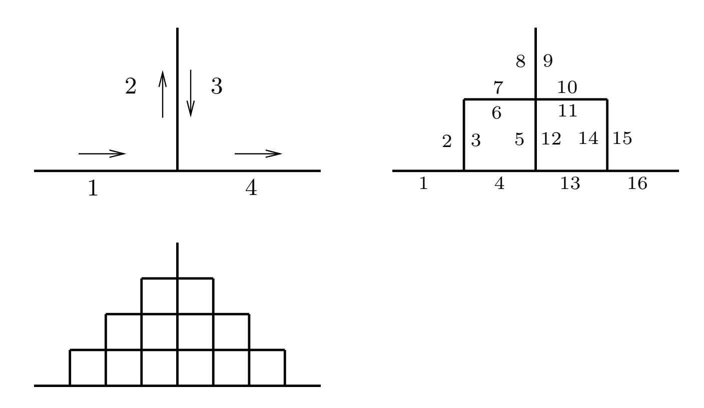


Figure 15. The first three steps of the von Koch curve when $\ell = 1/2$

16. Prove that the von Koch curve $t \mapsto \mathcal{K}^{\ell}(t)$, $1/4 < \ell \leq 1/2$ is continuous but nowhere differentiable.

[Hint: If $\mathcal{K}'(t)$ exists for some t, then

$$\lim_{n\to\infty} \frac{\mathcal{K}(u_n) - \mathcal{K}(v_n)}{u_n - v_n}$$

must exist, where $u_n \leq t \leq v_n$, and $u_n - v_n \to 0$. Choose $u_n = k/4^n$ and $v_n = (k+1)/4^n$.]

17. For a compact set E in $\mathbb{R}^d$, define $\#(\epsilon)$ to be the least number of balls of radius ϵ that cover E. Note that we always have $\#(\epsilon) = O(\epsilon^{-d})$ as $\epsilon \to 0$, and $\#(\epsilon) = O(1)$ if E is finite.

One defines the **covering dimension** of E, denoted by $\dim_C(E)$, as $\inf \beta$ such that $\#(\epsilon) = O(\epsilon^{-\beta})$, as $\epsilon \to 0$. Show that $\dim_C(E) = \dim_M(E)$, where $\dim_M$ is the Minkowski dimension discussed in Section 2.1, by proving the following inequalities for all $\delta > 0$:

(i) $m(E^\delta) \leq c\#(\delta)\delta^d$.

(ii) $\#(\delta) \leq c' m(E^\delta)\delta^{-d}$.

[Hint: To prove (ii), use Lemma 1.2 in Chapter 3 to find a collection of disjoint balls $B_1, B_2, \ldots, B_N$ of radius $\delta/3$, each centered at E, such that their "triples" $\tilde{B}_1, \tilde{B}_2, \ldots, \tilde{B}_N$ (of radius δ) cover E. Then $\#(\delta) \leq N$, while $Nm(B_j) = cN\delta^d \leq m(E^\delta)$, since the balls B_j are disjoint and are contained in E^δ.]

18. Let E be a compact set in $\mathbb{R}^d$.

(a) Prove that $\dim(E) \leq \dim_M(E)$, where dim and $\dim_M$ are the Hausdorff and Minkowski dimensions, respectively.

(b) However, prove that if $E = \{0, 1/\log 2, 1/\log 3, \ldots, 1/\log n, \ldots\}$, then $\dim_M E = 1$, yet $\dim E = 0$.

19. Show that there is a constant c_d, dependent only on the dimension d, such that whenever E is a compact set,

$$m(E^{2\delta}) \leq c_d m(E^\delta).$$

[Hint: Consider the maximal function f^*, with $f = \chi_{E^\delta}$, and take $c_d = 6^d$.]

20. Show that if F is the self-similar set considered in Theorem 2.12, then it has the same Minkowski dimension as Hausdorff dimension.
[Hint: Each F_k is the union of m^k balls of radius cr^k. In the converse direction one sees by Lemma 2.13 that if $\epsilon = r^k$, then each ball of radius ϵ can contain at most c' vertices of the kth generation. So it takes at least m^k/c' such balls to cover F.]

21. From the unit interval, remove the second and fourth quarters (open intervals). Repeat this process in the remaining two closed intervals, and so on. Let F be the limiting set, so that

$$F = \{x : \; x = \sum_{k=1}^{\infty} a_k/4^k \quad a_k = 0 \text{ or } 2\}.$$

Prove that $0 < m_{1/2}(F) < \infty$.

22. Suppose F is the self-similar set arising in Theorem 2.9.

(a) Show that if $m \leq 1/r^d$, then $m_d(F_i \cap F_j) = 0$ if $i \neq j$.

(b) However, if $m \geq 1/r^d$, prove that $F_i \cap F_j$ is not empty for some $i \neq j$.

(c) Prove that under the hypothesis of Theorem 2.12

$$m_\alpha(F_i \cap F_j) = 0, \qquad \text{with } \alpha = \log m/\log(1/r), \text{ whenever } i \neq j.$$

23. Suppose $S_1, \ldots, S_m$ are similarities with ratio r, $0 < r < 1$. For each set E, let

$$\tilde{S}(E) = S_1(E) \cup \cdots \cup S_m(E),$$

and suppose F denotes the unique non-empty compact set with $\tilde{S}(F) = F$.

(a) If $\overline{x} \in F$, show that the set of points $\{\tilde{S}^n(\overline{x})\}_{n=1}^{\infty}$ is dense in F.

(b) Show that F is **homogeneous** in the following sense: if $x_0 \in F$ and B is any open ball centered at x_0, then $F \cap B$ contains a set similar to F.

24. Suppose E is a Borel subset of $\mathbb{R}^d$ with $\dim E < 1$. Prove that E is totally disconnected, that is, any two distinct points in E belong to different connected components.

[Hint: Fix $x, y \in E$, and show that $f(t) = |t - x|$ is Lipschitz of order 1, and hence $\dim f(E) < 1$. Conclude that $f(E)$ has a dense complement in $\mathbb{R}$. Pick r in the complement of $f(E)$ so that $0 < r < f(y)$, and use the fact that $E = \{t \in E : |t - x| < r\} \cup \{t \in E : |t - x| > r\}$.]

25. Let $F(t)$ be an arbitrary non-negative measurable function on $\mathbb{R}$, and $\gamma \in S^{d-1}$. Then there exists a measurable set E in $\mathbb{R}^d$, such that $F(t) = m_{d-1}(E \cap \mathcal{P}_{t,\gamma})$.

26. Theorem 4.1 can be refined for $d \geq 4$ as follows.

Define $C^{k,\alpha}$ to be the class of functions $F(t)$ on $\mathbb{R}$ that are C^k and for which $F^{(k)}(t)$ satisfies a Lipschitz condition of exponent α.

If E has finite measure, then for a.e. $\gamma \in S^{d-1}$ the function $m(E \cap \mathcal{P}_{t,\gamma})$ is in $C^{k,\alpha}$ for $k = (d-3)/2$, $\alpha < 1/2$, if d is odd, $d \geq 3$; and for, $k = (d-4)/2$, $\alpha < 1$, if d is even, $d \geq 4$.

27. Show that the modification of the inequality (2) of Theorem 4.5 fails if we drop $\|f\|_{L^2(\mathbb{R}^d)}$ from the right-hand side.

[Hint: Consider $\mathcal{R}^*(f_\epsilon)$, with f_ϵ defined by $f_\epsilon(x) = (|x| + \epsilon)^{-d+\delta}$, for $|x| \leq 1$, with δ fixed, $0 < \delta < 1$, and $\epsilon \to 0$.]

28. Construct a compact set $E \subset \mathbb{R}^d$, $d \geq 3$, such that $m_d(E) = 0$, yet E contains translates of any segment of unit length in $\mathbb{R}^d$. (While particular examples of such sets can be easily obtained from the case $d = 2$, the determination of the least Hausdorff dimension among all such sets is an open problem.)

6 Problems

1. Carry out the construction below of two sets U and V so that

$$\dim U = \dim V = 0 \quad \text{but} \quad \dim(U \times V) \geq 1.$$

Let $I_1, \ldots, I_n, \ldots$ be given as follows:

- Each I_j is a finite sequence of consecutive positive integers; that is, for all j

$$I_j = \{n \in \mathbb{N} : A_j \leq n \leq B_j\} \quad \text{for some given } A_j \text{ and } B_j.$$

- For each j, I_{j+1} is to the right of I_j; that is, $A_{j+1} > B_j$.

Let $U \subset [0,1]$ consist of all x which when written dyadically $x = .a_1 a_2 \cdots a_n \cdots$ have the property that $a_n = 0$ whenever $n \in \bigcup_j I_j$. Assume also that A_j and B_j tend to infinity (as $j \to \infty$) rapidly enough, say $B_j/A_j \to \infty$ and $A_{j+1}/B_j \to \infty$.

Also, let J_j be the complementary blocks of integers, that is,

$$J_j = \{n \in \mathbb{N} : B_j < n < A_{j+1}\}.$$

Let $V \subset [0,1]$ consist of those $x = .a_1 a_2 \cdots a_n \cdots$ with $a_n = 0$ if $n \in \bigcup_j J_j$.

Prove that U and V have the desired property.

2.* The iso-diametric inequality states the following: If E is a bounded subset of $\mathbb{R}^d$ and diam $E = \sup\{|x - y| : x, y \in E\}$, then

$$m(E) \leq v_d \left(\frac{\text{diam } E}{2} \right)^d,$$

where v_d denotes the volume of the unit ball in $\mathbb{R}^d$. In other words, among sets of a given diameter, the ball has maximum volume. Clearly, it suffices to prove the inequality for $\overline{E}$ instead of E, so we can assume that E is compact.

(a) Prove the inequality in the special case when E is symmetric, that is, $-x \in E$ whenever $x \in E$.

In general, one reduces to the symmetric case by using a technique called Steiner symmetrization. If e is a unit vector in $\mathbb{R}^d$, and $\mathcal{P}$ is a plane perpendicular to e, the Steiner symmetrization of E with respect to E is defined by

$$S(E, e) = \{x + te : x \in \mathcal{P},\ |t| \leq \frac{1}{2} L(E; e; x)\},$$

where $L(E; e; x) = m\left(\{t \in \mathbb{R} :\ x + t \cdot e \in E\}\right)$, and m denotes the Lebesgue measure. Note that $x + te \in S(E, e)$ if and only if $x - te \in S(E, e)$.

(b) Prove that $S(E, e)$ is a bounded measurable subset of $\mathbb{R}^d$ that satisfies $m(S(E, e)) = m(E)$.

[Hint: Use Fubini's theorem.]

(c) Show that diam $S(E, e) \leq$ diam E.

(d) If ρ is a rotation that leaves E and $\mathcal{P}$ invariant, show that $\rho S(E, e) = S(E, e)$.

(e) Finally, consider the standard basis $\{e_1, \ldots, e_d\}$ of $\mathbb{R}^d$. Let $E_0 = E$, $E_1 = S(E_0, e_1)$, $E_2 = S(E_1, e_2)$, and so on. Use the fact that E_d is symmetric to prove the iso-diametric inequality.

(f) Use the iso-diametric inequality to show that $m(E) = \frac{v_d}{2^d} m_d(E)$ for any Borel set E in $\mathbb{R}^d$.

3. Suppose S is a similarity.

(a) Show that S maps a line segment to a line segment.

(b) Show that if L_1 and L_2 are two segments that make an angle α, then $S(L_1)$ and $S(L_2)$ make an angle α or $-\alpha$.

(c) Show that every similarity is a composition of a translation, a rotation (possibly improper), and a dilation.

4.* The following gives a generalization of the construction of the Cantor-Lebesgue function.

Let F be the compact set in Theorem 2.9 defined in terms of m similarities $S_1, S_2, \ldots, S_m$ with ratio $0 < r < 1$. There exists a unique Borel measure μ supported on F such that $\mu(F) = 1$ and

$$\mu(E) = \frac{1}{m} \sum_{j=1}^{m} \mu(S_j^{-1}(E)) \qquad \text{for any Borel set } E.$$

In the case when F is the Cantor set, the Cantor-Lebesgue function is $\mu([0, x])$.

5. Prove a theorem of Hausdorff: Any compact subset K of $\mathbb{R}^d$ is a continuous image of the Cantor set $\mathcal{C}$.

[Hint: Cover K by 2^{n_1} (some n_1) open balls of radius 1, say $B_1, \ldots, B_\ell$ (with possible repetitions). Let $K_{j_1} = K \cap \overline{B_{j_1}}$ and cover each K_{j_1} with 2^{n_2} balls of radius 1/2 to obtain compact sets K_{j_1,j_2}, and so on. Express $t \in \mathcal{C}$ as a ternary expansion, and assign to t a unique point in K defined by the intersection $K_{j_1} \cap K_{j_1,j_2} \cap \cdots$ for appropriate $j_1, j_2, \ldots$. To prove continuity, observe that if two points in the Cantor set are close, then their ternary expansions agree to high order.]

6. A compact subset K of $\mathbb{R}^d$ is **uniformly locally connected** if given $\epsilon > 0$ there exists $\delta > 0$ so that whenever $x, y \in K$ and $|x - y| < \delta$, there is a continuous curve γ in K joining x to y, such that $\gamma \subset B_\epsilon(x)$ and $\gamma \subset B_\epsilon(y)$.

Using the previous problem, one can show that a compact subset K of $\mathbb{R}^d$ is the continuous image of the unit interval $[0, 1]$ if and only if K is uniformly locally connected.

7. Formulate and prove a generalization of Theorem 3.5 to the effect that once appropriate sets of measure zero are removed, there is a measure-preserving isomorphism of the unit interval in $\mathbb{R}$ and the unit cube in $\mathbb{R}^d$.

8.* There exists a *simple* continuous curve in the plane of positive two-dimensional measure.

9. Let E be a compact set in $\mathbb{R}^{d-1}$. Show that $\dim(E \times I) = \dim(E) + 1$, where I is the unit interval in $\mathbb{R}$.

10.* Let $\mathcal{C}_\xi$ be the Cantor set considered in Exercises 8 and 11. If $\xi < 1/2$, then $\mathcal{C}_\xi + \lambda\mathcal{C}_\xi$ has positive Lebesgue measure for almost every λ.

Notes and References

There are several excellent books that cover many of the subjects treated here. Among these texts are Riesz and Nagy [27], Wheeden and Zygmund [33], Folland [13], and Bruckner *et al.* [4].

Introduction

The citation is a translation of a passage in a letter from Hermite to Stieltjes [18].

Chapter 1

The citation is a translation from the French of a passage in [3].

We refer to Devlin [7] for more details about the axiom of choice, Hausdorff maximal principle, and well-ordering principle.

See the expository paper of Gardner [14] for a survey of results regarding the Brunn-Minkowski inequality.

Chapter 2

The citation is a passage from the preface to the first edition of Lebesgue's book on integration [20].

Devlin [7] contains a discussion of the continuum hypothesis.

Chapter 3

The citation is from Hardy and Littlewood's paper [15].

Hardy and Littlewood proved Theorem 1.1 in the one-dimensional case by using the idea of rearrangements. The present form is due to Wiener.

Our treatment of the isoperimetric inequality is based on Federer [11]. This work also contains significant generalizations and much additional material on geometric measure theory.

A proof of the Besicovitch covering in the lemma in Problem 3^* is in Mattila [22].

For an account of functions of bounded variations in $\mathbb{R}^d$, see Evans and Gariepy [8].

An outline of the proof of Problem 7 (b)* can be found at the end of Chapter 5 in Book I.

The result in part (b) of Problem 8^* is a theorem of S. Saks, and its proof as a consequence of part (a) can be found in Stein [31].

Chapter 4

The citation is translated from the introduction of Plancherel's article [25].

An account of the theory of almost periodic functions which is touched upon in Problem 2^* can be found in Bohr [2].

The results in Problems 4^* and 5^* are in Zygmund [35], in Chapters V and VII, respectively.

Consult Birkhoff and Rota [1] for more on Sturm-Liouville systems, Legendre polynomials, and Hermite functions.

Chapter 5

See Courant [6] for an account of the Dirichlet principle and some of its applications. The solution of the Dirichlet problem for general domains in $\mathbb{R}^2$ and the related notion of logarithmic capacity of sets are treated in Ransford [26]. Folland [12] contains another solution to the Dirichlet problem (valid in $\mathbb{R}^d$, $d \geq 2$) by methods which do not use the Dirichlet principle.

The result regarding the existence of the conformal mapping stated in Problem 3^* is in Chapter VII of Zygmund [35].

Chapter 6

The citation is a translation from the German of a passage in C. Carathéodory [5].

Petersen [24] gives a systematic presentation of ergodic theory, including a proof of the theorem in Problem 7^*.

The facts about spherical harmonics needed in Problem 4^* can be found in Chapter 4 in Stein and Weiss [32].

We refer to Hardy and Wright [16] for an introduction to continued fractions. Their connection to ergodic theory is discussed in Ryll-Nardzewski [28].

Chapter 7

The citation is a translation from the German of a passage in Hausdorff's article [17], while Mandelbrot's citation is from his book [21].

Mandelbrot's book also contains many interesting examples of fractals arising in a variety of different settings, including a discussion of Richardson's work on the length of coastlines. (See in particular Chapter 5.)

Falconer [10] gives a systematic treatment of fractals and Hausdorff dimension.

We refer to Sagan [29] for further details on space-filling curves, including the construction of a curve arising in Problem 8^*.

The monograph of Falconer [10] also contains an alternate construction of the Besicovitch set, as well as the fact that such sets must necessarily have dimension two. The particular Besicovitch set described in the text appears in Kahane [19], but the fact that it has measure zero required further ideas which are contained, for instance, in Peres *et al.* [30].

Regularity of sets in $\mathbb{R}^d$, $d \geq 3$, and the estimates for the maximal function associated to the Radon transform are in Falconer [9], and Oberlin and Stein [23].

The theory of Besicovitch sets in higher dimensions, as well as a number of interesting related topics can be found in the survey of Wolff [34].

Bibliography

[1] G. Birkhoff and G. C. Rota. *Ordinary differential equations.* Wiley, New York, 1989.

[2] H. A. Bohr. *Almost periodic functions.* Chelsea Publishing Company, New York, 1947.

[3] E. Borel. *Leçons sur la théorie des fonctions.* Gauthiers-Villars, Paris, 1898.

[4] J. B. Bruckner, A. M. Bruckner, and B. S. Thomson. *Real Analysis.* Prentice Hall, Upper Saddle River, NJ, 1997.

[5] C. Carathéodory. *Vorlesungen über reelle Funktionen.* Leipzig, Berlin, B. G. Teubner, Leipzig and Berlin, 1918.

[6] R. Courant. *Dirichlet's principle, conformal mappings, and minimal surfaces.* Interscience Publishers, New York, 1950.

[7] K. J. Devlin. *The joy of sets: fundamentals of contemporary set theory.* Springer-Verlag, New York, 1997.

[8] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* CRC Press, Boca Raton, 1992.

[9] K. J. Falconer. Continuity properties of k-plane integrals and Besicovitch sets. *Math. Proc. Cambridge Philos. Soc*, 87:221–226, 1980.

[10] K. J. Falconer. *The geometry of fractal sets.* Cambridge University Press, 1985.

[11] H. Federer. *Geometric measure theory.* Springer, Berlin and New York, 1996.

[12] G. B. Folland. *Introduction to partial differential equations.* Princeton University Press, Princeton, NJ, second edition, 1995.

[13] G. B. Folland. *Real Analysis: modern techniques and their applications.* Wiley, New York, second edition, 1999.

[14] R. J. Gardner. The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc*, 39:355–405, 2002.

[15] G. H. Hardy and J. E. Littlewood. A maximal theorem with function theoretic applications. *Acta. Math*, 54:81–116, 1930.

[16] G. H. Hardy and E. M. Wright. *An introduction to the Theory of Numbers*. Oxford University Press, London, fifth edition, 1979.

[17] F. Hausdorff. Dimension und äusseres Mass. *Math. Annalen*, 79:157–179, 1919.

[18] C. Hermite. *Correspondance d'Hermite et de Stieltjes*. Gauthier-Villars, Paris, 1905. Edited by B. Baillaud and H. Bourget.

[19] J. P. Kahane. Trois notes sur les ensembles parfaits linéaires. *Enseignement Math.*, 15:185–192, 1969.

[20] H. Lebesgue. *Leçons sur l'integration et la recherche des fonctions primitives*. Gauthier-Villars, Paris, 1904. Preface to the first edition.

[21] B. B. Mandelbrot. *The fractal geometry of nature*. W. H. Freeman, San Francisco, 1982.

[22] P. Mattila. *Geometry of sets and measures in Euclidean spaces*. Cambridge University Press, Cambridge, 1995.

[23] D. M. Oberlin and E. M. Stein. Mapping properties of the Radon transform. *Indiana Univ. Math. J*, 31:641–650, 1982.

[24] K. E. Petersen. *Ergodic theory*. Cambridge University Press, Cambridge, 1983.

[25] M. Plancherel. La théorie des équations intégrales. *L'Enseignement math.*, 14e Année:89–107, 1912.

[26] T. J. Ransford. *Potential theory in the complex plane*. London Mathematical Society student texts, 28. Cambridge, New York: Press Syndicate of the University of Cambridge, 1995.

[27] F. Riesz and B. Sz.-Nagy. *Functional Analysis*. New York, Ungar, 1955.

[28] C. Ryll-Nardzewski. On the ergodic theorem. ii. Ergodic theory of continued fractions. *Studia Math.*, 12:74–79, 1951.

[29] H. Sagan. *Space-filling curves*. Universitext. Springer-Verlag, New York, 1994.

[30] Y. Peres, K. Simon, and B. Solomyak. Fractals with positive length and zero Buffon needle probability. *Amer. Math. Monthly*, 110:314–325, 2003.

[31] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton University Press, Princeton, NJ, 1993.

[32] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, Princeton, NJ, 1971.

[33] R. L. Wheeden and A. Zygmund. *Measure and integral: an introduction to real analysis.* Marcel Dekker, New York, 1977.

[34] T. Wolff. Recent work connected with the Kakeya problem. *Prospects in Mathematics,Princeton, NJ*, 31:129–162, 1996. Amer. Math. Soc., Providence, RI, 1999.

[35] A. Zygmund. *Trigonometric Series*, volume I and II. Cambridge University Press, Cambridge, second edition, 1959. Reprinted 1993.

Symbol Glossary

The page numbers on the right indicate the first time the symbol or notation is defined or used. As usual, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ denote the integers, the rationals, the reals, and the complex numbers respectively.

$\lvert x\rvert$	(Euclidean) Norm of x	2
E^c, $E - F$	Complements and relative complements of sets	2
$d(E, F)$	Distance between two sets	2
$B_r(x), \overline{B_r(x)}$	Open and closed balls	2
$\overline{E}$, ∂E	Closure and boundary of E, respectively	3
$\lvert R\rvert$	Volume of the rectangle R	3
$O(\cdots)$	O notation	12
$\mathcal{C}$, $\mathcal{C}_\xi$, $\hat{\mathcal{C}}$	Cantor sets	9, 38
$m_*(E)$	Exterior (Lebesgue) measure of the set E	10
$E_k \nearrow E$, $E_k \searrow E$	Increasing and decreasing sequences of sets	20
$E \triangle F$	Symmetric difference of E and F	21
$E_h = E + h$	Translation by h of the set E	22
$\mathcal{B}_{\mathbb{R}^d}$	Borel σ-algebra on $\mathbb{R}^d$	23
G_δ, F_σ	Sets of type G_δ or F_σ	23
$\mathcal{N}$	Non-measurable set	24
a.e.	Almost everywhere	30
$f^+(x)$, $f^-(x)$	Positive and negative parts of f	31, 64
$A + B$	Sum of two sets	35
v_d	Volume of the unit ball in $\mathbb{R}^d$	39
$\text{supp}(f)$	Support of the function f	53
$f_k \nearrow f$, $f_k \searrow f$	Increasing and decreasing sequences of functions	62
f_h	Translation by h of the function f	73
$L^1(\mathbb{R}^d)$, $L^1_{\text{loc}}(\mathbb{R}^d)$	Integrable and locally integrable functions	69, 105
$f * g$	Convolution of f and g	74
f^y, f_x, E^y, E_x	Slices of the function f and set E	75
$\hat{f}$, $\mathcal{F}(f)$	Fourier transform of f	87, 208
f^*	Maximal functions of f	100, 296
$L(\gamma)$	Length of the (rectifiable) curve γ	115
T_F, P_F, N_F	Total, positive, and negative variations of F	117, 118
$L(A, B)$	Length of a curve between $t = A$ and $t = B$	120
$D^+(F), \ldots, D_-(F)$	Dini numbers of F	123

$\mathcal{M}(K)$	Minkowski content of K	138
$\Omega_+(\delta)$, $\Omega_-(\delta)$	Outer and inner set of Ω	143
$L^2(\mathbb{R}^d)$	Square integrable functions	156
$\ell^2(\mathbb{Z})$, $\ell^2(\mathbb{N})$	Square summable sequences	163
$\mathcal{H}$	Hilbert space	161
$f \perp g$	Orthogonal elements	164
$\mathbb{D}$	Unit disc	173
$H^2(\mathbb{D})$, $H^2(\mathbb{R}^2_+)$	Hardy spaces	174, 213
$\mathcal{S}^\perp$	Orthogonal complement of $\mathcal{S}$	177
$A \oplus B$	Direct sum of A and B	177
$P_\mathcal{S}$	Orthogonal projection onto $\mathcal{S}$	178
T^*, L^*	Adjoint of operators	183, 222
$\mathcal{S}(\mathbb{R}^d)$	Schwartz space	208
$C_0^\infty(\Omega)$	Smooth functions with compact support in Ω	222
$C^n(\Omega)$, $C^n(\overline{\Omega})$	Functions with n continuous derivatives on Ω and $\overline{\Omega}$	223
$\triangle u$	Laplacian of u	230
$(X, \mathcal{M}, \mu)$, (X, μ)	Measure space	263
μ, μ_*, μ_0	Measure, exterior measure, premeasure	263, 264, 270
$\mu_1 \times \mu_2$	Product measure	276
S^{d-1}	Unit sphere in $\mathbb{R}^d$	279
σ, $d\sigma(\gamma)$	Surface measure on the sphere	280
dF	Lebesgue-Stieltjes measure	282
$\vert\nu\vert$, ν^+, ν^-	Total, positive, and negative variations of ν	286, 287
$\nu \perp \mu$	Mutually singular measures	288
$\nu \ll \mu$	Absolutely continuous measures	289
$\sigma(S)$	Spectrum of S	311
$m_\alpha^*(E)$	Exterior α-dimensional Hausdorff measure	325
diam S	Diameter of S	325
$\dim E$	Hausdorff dimension of E	329
$\mathcal{S}$	Sierpinski triangle	334
$A \approx B$	A comparable to B	335
$\mathcal{K}$, $\mathcal{K}^\ell$	Von Koch curves	338, 340
$\text{dist}(A, B)$	Hausdorff distance	345
$\mathcal{P}(t)$	Peano mapping	349
$\mathcal{P}_{t,\gamma}$	Hyperplane	360
$\mathcal{R}(f)$, $\mathcal{R}_\delta(f)$	Radon transform	363, 368
$\mathcal{R}^*(f)$, $\mathcal{R}_\delta^*(f)$	Maximal Radon transform	363, 368

Index

Relevant items that also arose in Book I or Book II are listed in this index, preceeded by the numerals I or II, respectively.